

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 06-11-03
Searcher: Beverly C4994
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 23
Number of Searches: _____
Number of Databases: 1

Search Site

_____ STIC
_____ CM-1
_____ Pre-S

Type of Search

_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors

_____ IG Suite
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
_____ Other CGN

THIS PAGE BLANK (USPTO)

Result No.	Query			ID	Description
	Score	Match	Length		
1	225	100.0	225	10	US-09-765-231A-58
2	38.2	17.0	640681	10	Sequence 58, Appl
3	37.4	16.6	2268	9	Sequence 1, Appl
4	36.4	16.2	1479	10	Sequence 2, Appl
5	36.4	16.2	1475	10	Sequence 772, Appl
6	35.8	15.9	885	9	Sequence 1661, Appl
7	35.2	15.6	2324	10	Sequence 4260, Appl
8	34.6	15.4	1455	10	Sequence 4454, Appl
9	34.6	15.4	1455	10	Sequence 104, Appl
10	34.6	15.4	1747	9	Sequence 106, Appl
11	34.4	15.3	379	9	Sequence 10020, Appl
12	34.4	15.3	1132	9	Sequence 12188, Appl
13	34.4	15.3	2749	9	Sequence 98, Appl
14	34.4	15.3	3161	10	Sequence 28, Appl
15	34.4	15.3	640681	10	Sequence 577, Appl
16	34.2	15.2	824	9	Sequence 1, Appl
17	34.2	15.2	30801	9	Sequence 13, Appl
18	34.2	15.2	90650	9	Sequence 33, Appl
19	34	15.1	584	9	Sequence 80, Appl
				9	Sequence 2363, Appl

RESULT 2

US-09-790-988-1/c

; Sequence 1, Application US/09790988

; Patent No. US20020127687A1

; GENERAL INFORMATION:

; APPLICANT: SHIGENOBU, SHUJI

; APPLICANT: WATANABE, HIDEMI

; APPLICANT: HATTORI, MASAHIRA

; APPLICANT: SAKAKI, YOSHIYUKI

; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS

; FILE REFERENCE: 081356/0159

; CURRENT APPLICATION NUMBER: US/09/790,988

; CURRENT FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: JP2000-107160

; PRIOR FILING DATE: 2000-04-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 640681

; TYPE: DNA

; ORGANISM: Buchnera sp.

; US-09-790-988-1

Query Match 17.0%; Score 38.2; DB 10; Length 640681;

Best Local Similarity 55.7%; Pred. No. 19;

Matches 73; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY

32 AATAAATTAGAGGCTCCATGATATGCTATATGTTTACCTTCAGAGAATATTAG 91

Db

198302 AATCAAAATTAAATAAAGCTAGTATATTTCAATTATGACTTTTCAATAATTTTATTATTAAAG 198243

QY

92 TTTCACTCAGGTTTTTCAAGCTAGCTGTCCTCCCAAAACGAAACAAACAAAAAAC 151

Db

198242 TAATAATTGTTATTATATAATAACTTTTCATCTAAATAAACAATCAAAAAA 198183

QY

152 AACCTTTTAA 162

Db

198182 AAGATTITTTA 198172

RESULT 3

US-10-072-130-2/c

; Sequence 2, Application US/10072130

; Patent No. US20020173022A1

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; Goli, Surya K.

; Lal, Preeti

; Corley, Neil C.

; Zhang, Hong

; TITLE OF INVENTION: NEW PROTEIN PHOSPHATASE

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/072,130

; FILING DATE: 05-Feb-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/873,093

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0319 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2268 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: 13177

CLONE: THPIPLB01

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-072-130-2

Query Match 16.6%; Score 37.4; DB 9; Length 2268;

Best Local Similarity 53.0%; Pred. No. 2.9;

Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY

29 TTGAATAAATTATGAGGCTCCATGATATGCTATATGTTTACCTTCAGAGAATATT 88

Db

2195 TTGACATACATTTTGATGTGTGTAGTGTAGCTATACAGAAATATACAAACCAAGCTGTAT 2136

QY

89 TAGTTTCACTCAGGTTTTTCAAGCTAGCTGCTCCCAAAACGAAACAAACAAACAAA 148

Db

2135 GAATAATACATAGTTTTTCAAAAGTTATGTTTCATAAAGATACAGAGAAAGGAAAT 2076

QY

149 AACAACTTTTAAAGATGATGGCTACTCA 179

Db

2075 AGCCACACATCCAATATCTCACAACCTTTCAA 2045

RESULT 4

US-09-770-791-772/c

; Sequence 772, Application US/09770791

; Patent No. US20020062014A1

; GENERAL INFORMATION:

; APPLICANT: Goriach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kricker, Meja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; TITLE OF INVENTION: thaliana

; FILE REFERENCE: 2029 (PARA-018PRV)

; CURRENT APPLICATION NUMBER: US/09/770,791

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,480

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 772

; LENGTH: 345

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; US-09-770-791-772

Query Match 16.2%; Score 36.4; DB 10; Length 345;
Best Local Similarity 53.5%; Pred. No. 2.5;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 12 TGTTCAGGCATAAATTTGAAATAATATGAGGCTCCATGATATGCTATATTTGGTTT 71
DB 155 TGTATTACCTAGACATAAATATGATTATGATAAACTAATTGTAATTACTGCATAT 96

QY 72 ACCTTCAGAGAATATTTAGTTTCATCAGGTTTTTCAAAGCTACGCTGTCCTCCCAAAA 131
DB 95 GAAATATATAATCTTTTGTGTTTAACTCAAGATGTAATAATTTATGATCTTCACATAAAA 36

QY 132 ACCAACAACAACAAAAACAA 153
DB 35 AAAAAAAAAAAAAAAAAAAAAA 14

RESULT 5
US-09-880-107-1661
Sequence 1661, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1661
LENGTH: 1479
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 D45288
US-09-880-107-1661

Query Match 16.2%; Score 36.4; DB 10; Length 1479;
Best Local Similarity 61.7%; Pred. No. 4.5;
Matches 58; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

DB 128 AAAAAACAACAACAACAACAACCTTTTAAAGATTGATGGCTACTCATTTGATCT 187
DB 834 AGAAAAGAAACAACAACAACAACCTTTTAAATTTACTTGAACCTCAACAGATCT 893

QY 188 GCCTCTCTGCTCAATCAATTAGGAATTTTTTT 221
DB 894 CCCTGCGTACTCCCTTTTCCAGGAATTTACTT 927

RESULT 6
US-10-198-846-4260/c
Sequence 4260, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220

Query Match 15.9%; Score 35.8; DB 9; Length 885;
Best Local Similarity 53.8%; Pred. No. 5.3;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 27 ATTGAAATAAATTTATGAGGCTCCATGATATGCTATATTTGGTTTACCTTCAGAAATA 86
DB 169 AATTAGCTANATTTATCTTNTCAATCAATCTTACTTCAGAGTAAATAAAAAAGT 110

QY 87 TTTAGTTTCACCTCAGGTTTTTCAAAGCTACGCTGTCCTCCCAAAAAACGAAACAA 146
DB 109 TATGCAATACATAAATAATGCAAAAAAGGTATNCCAAAAAATAAAAAAAAAA 50

QY 147 AAACAACCT 156
DB 49 AAAAGTACCT 40

RESULT 7
US-09-983-965-4454/c
Sequence 4454, Application US/09983965
Patent No. US20020137160A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 37-21(10297)C
CURRENT APPLICATION NUMBER: US/09/983,965
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 4454
LENGTH: 234
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
OTHER INFORMATION: Clone ID: 05-LIB34-080-Q1-E1-B9
US-09-983-965-4454

Query Match 15.6%; Score 35.2; DB 10; Length 234;
Best Local Similarity 58.7%; Pred. No. 4.4;
Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 4 TCGTAAGTTGTTTCAGGCATAAATTTGAAATAAATTTATGAGCTCCATGATGCTATA 63

Db 159 TGATGCTTTAGACAGATAGAGTATGAAAAAATTTCTGATAACATGATATATAAAA 100
QY 64 TTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTT 107
Db 99 ATGTTATTACATGCGAAGAAATAGTACCCCTCACAGAGTATT 56

RESULT 8

US-09-917-265-104
; Sequence 104, Application US/09917265
; Patent No. US20020052030A1
; GENERAL INFORMATION:
; APPLICANT: Wonderling, Ramani S.
; APPLICANT: Boroughs, Karen L.
; TITLE OF INVENTION: CANINE AND FELINE PROTEINS, NUCLEIC ACID MOLECULES AND USES THERE
; FILE REFERENCE: IM-5
; CURRENT APPLICATION NUMBER: US/09/917,265
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/223,016
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
; TYPE: DNA
; LENGTH: 1455
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)..(897)
; OTHER INFORMATION:
US-09-917-265-104

Query Match 15.4%; Score 34.6; DB 10; Length 1455;
Best Local Similarity 51.6%; Pred. No. 13; Mismatches 0; Gaps 0;
Matches 79; Conservative 0

QY 1 TGATGCTAGCTTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60
Db 1298 TTATATAATTTTGTCTCATGAAGCATGTGAATTAATTTATTTATTTATGTTATTT 1357
QY 61 ATATTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTAGCGTG 120
Db 1358 TATTAAAGTATTTATTATCATCAAGTGGATTTGGGATATCTTATGTTCTAAATAAATGAT 1417
QY 121 TCCCCCAAAACGAAACAAACAAACAAACAA 153
Db 1418 TGAGTAGAAAAAATAAATAAATAAATAAATAA 1450

RESULT 9

US-09-917-265-106/c
; Sequence 106, Application US/09917265
; Patent No. US20020052030A1
; GENERAL INFORMATION:
; APPLICANT: Wonderling, Ramani S.
; APPLICANT: Boroughs, Karen L.
; TITLE OF INVENTION: CANINE AND FELINE PROTEINS, NUCLEIC ACID MOLECULES AND USES THERE
; FILE REFERENCE: IM-5
; CURRENT APPLICATION NUMBER: US/09/917,265
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/223,016
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-917-265-106

Query Match 15.4%; Score 34.6; DB 10; Length 1455;
Best Local Similarity 51.6%; Pred. No. 13;
Matches 79; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 TGATGCTAGCTTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60
Db 158 TTATATAATTTTGTCTCATGAAGCATGTGAATTAATTTATTTATTTATTTATTT 99
QY 61 ATATTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTAGCGTG 120
Db 98 TATTAAAGTATTTATTATCATCAAGTGGATTTGGGATATCTTATGTTCTAAATAAATGAT 39
QY 121 TCCCCCAAAACGAAACAAACAAACAAACAA 153
Db 38 TGAGTAGAAAAAATAAATAAATAAATAAATAA 6

RESULT 10

US-10-198-846-10020/c
; Sequence 10020, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10020
; LENGTH: 1747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 1747
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10020

Query Match 15.4%; Score 34.6; DB 9; Length 1747;
Best Local Similarity 55.4%; Pred. No. 15;
Matches 57; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 30 TGAATAAATATGAGGCTCCATGATATGCTATATTTGTTTACCTTCAGAGAAATATT 89
Db 1390 TAAATAATATATATCAGGCCCATCTCTTTTTTTTTTTTTTTTTTTTAAAGATGGAGTCTC 1331
QY 90 AGTTTCACTCAGGTTTTTCAAAGCTAGCGTCCCCCAAAACGAAACAAACAAACAA 149
Db 1330 ACTGTGTACCCAGGCTGGAGTGAGACTCTGTATCGAAGAAAAAATAAAGAGGAA 1271
QY 150 A 150
Db 1270 A 1270

RESULT 11

US-10-198-846-12188/c
; Sequence 12188, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12188
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 378..379
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12188

Query Match 15.3%; Score 34.4; DB 9; Length 379;
Best Local Similarity 54.8%; Pred. No. 8.7;
Matches 58; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
27 ATTTGAAATAAATATGAGGCTCCATGATATGCTATATTGTTTACCTTCAGAAGAATA 86
126 AATTAGCTATATTATCTCTCAATCAAAATCTCTTCAGAGGTAAGTAAAAAAGAAGT 67
QY 87 TTTAGTTTCACTCAGGTTTTTCAAAGCTAGCTGTCCCCCAAAAAACGAAACAAACAA 146
Db 66 TATAGCAATACATAAAATGGAACAAAAGGTATCCCAAAAAAAGGTAAGTAAAAA 7
QY 147 AAAA 150
Db 6 AAAA 3

RESULT 12
US-09-986-480-98
; Sequence 98, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; FILE REFERENCE: PS500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (153)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (260)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-986-480-98

Query Match 15.3%; Score 34.4; DB 9; Length 1132;
Best Local Similarity 54.8%; Pred. No. 14;
Matches 58; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 30 TGAATAAATATGAGGCTCCATGATATGCTATATTGTTTACCTTCAGAAGAATAATT 89
Db 1009 TGAAGAATTTTGTATTATACAGAAGAGCCAGATGAATATTGTTTAAAGTAGAGCATTT 1068
QY 90 ACTTTCACTCAGGTTTTTCAAAGCTAGCTGTCCCAAAAAACGAAACAAACAAACAA 149
Db 1069 CCTTTTATTAAATAAATTAACAAATGTTTACATCCAAAAAAGGTAAGTAAAAA 1128

QY 150 ACAA 153
Db 1129 TCGA 1132
RESULT 13
US-09-813-153-29
; Sequence 29, Application US/09813153
; Publication No. US20030045459A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 67 Human secreted proteins
; FILE REFERENCE: P2023
; CURRENT APPLICATION NUMBER: US/09/813,153
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/363,044
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,167
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,162
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,161
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,170
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 2749
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-813-153-29

Query Match 15.3%; Score 34.4; DB 9; Length 2749;
Best Local Similarity 51.3%; Pred. No. 20;
Matches 77; Conservative 1; Mismatches 72; Indels 0; Gaps 0;
QY 4 TGGTAAGTTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATATGCTATA 63
Db 2578 TGATATGATATTTTGTGTAAGTTTTTTGTAAATAAATTTTACAAATGTTTATTTGAAT 2637
QY 64 TTGGTTTTACCTTCAGAGAATATTAGTTTTCACCTCAGGTTTTTCAAAGCTAGCTGTCC 123
Db 2638 GATTTTTTAAATGCTGTGAATCTATATTGTTGTTTATATATAAATTCATTGGCCA 2697
QY 124 CCCAAAAACGAAACAAACAAAAACAA 153
Db 2698 AAAAAAAGGTAAGTAAAAAAGGTAAGTAAAAAAGGTAAGTAAAAAAGGTAAGTAAAAA 2727

RESULT 14
US-09-925-301-577
; Sequence 577, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12

Result No.	Score	Query Match	Length	DB	ID	Description
1	196	87.1	341	9	AA502552	AA502552 ng62e06.s
2	117	52.0	324	9	AI347782	AI347782 qh61a07.x
3	117	52.0	362	9	AA682512	AA682512 zi19a01.s
4	48.8	21.7	928	17	CNS00DKY	AL071865 Drosophi
5	43	19.1	773	17	AQ781761	AQ781761 HS 3122
6	41.8	18.6	964	17	CNS058MA	AL326107 Tetraodon

RESULT 6

CNS058MA/c
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T3 end of clone
002B22 of library A from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION
AL326107.1 GI:8219696
VERSION
GSS: genome survey sequence.
KEYWORDS
Tetraodon nigroviridis.
SOURCE
Tetraodon nigroviridis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 964)

REFERENCE
AUTHORS
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 964)

TITLE
JOURNAL
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
3 (bases 1 to 964)

REFERENCE
AUTHORS
Direct Submission
Genoscope.
Tetraodon nigroviridis
Submitted (12-APR-2000)
JOURNAL
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1. 964
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="002B22"
/clone_lib="A"
/notes="Genoscope sequence ID : COA002DALLA1-end : T3"

BASE COUNT
290 a 194 c 170 g 269 t 41 others

ORIGIN
Query Match 18.6%; Score 41.8; DB 17; Length 964;
Best Local Similarity 52.4%; Pred. No. 34;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 4 TGGTAGTTCTTCAGGCATAAATTTGAATAAATATGAGCTCCATCATATGCTATA 63
|||||
Db 821 TGGTAGTTNCCGGTTGTGACCAATTTTAGCCGTACAAATATGGGATTTCTATGTTAA 762
|||||
QY 64 TTGGTTTACCTTCAGAGAATATTTAGTTTCTACTCAGGTTTTTCAAAGCTAGCGTGCC 123
|||||
Db 761 NATTTCTCATTAACACATATGTTTATGTCATGATGATGTCNAAAATTAGTTGTAT 702
|||||
QY 124 CCAAAAAACGAAACAAAAACAAAAACAACTTTTAAAGATTGATG 171
|||||
Db 701 TTAATAAAAAACAAAAAATCAAGACAAGTTCTTTTCAGATTCTG 654
|||||

RESULT 7
CNS00D90/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC27P19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL060402
VERSION
AL060402.1 GI:4947764
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster.

ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscophora;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)

REFERENCE
AUTHORS
Direct Submission
Genoscope.
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR27P19"
/clone_lib="RPCI-98"
/note="end : TET3"

BASE COUNT
380 a 163 c 143 g 199 t 216 others

ORIGIN
Query Match 18.3%; Score 41.2; DB 17; Length 1101;
Best Local Similarity 34.5%; Pred. No. 41;
Matches 58; Conservative 42; Mismatches 68; Indels 0; Gaps 0;
QY 18 AGCGTAAATTTGAATAAATATGAGCTCCATCATATGCTATGTTTACCTTC 77
|||||
Db 1075 GAAATAAARADHWTATATATGRTTWTWKTWTWTWTWTWTWTWTWTWTWT 1016
|||||
QY 78 AGAAGATATTTAGTTTCTACTCAGGTTTTTCAAAGCTAGCTGCTCCCAAAACGAAA 137
|||||
Db 1015 AAAAAAAMATTTCTTAAATTTTWTWYTCYCWACAMCTCCCTTMCACMAWMC 956
|||||
QY 138 CAAAAACAAAAACAACTTTTAAAGATTGATGCTACTATTGAT 185
|||||
Db 955 CAAAAYAAAAATWMAKTTTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 908
|||||

RESULT 8
AW504318
LOCUS
DEFINITION
UI-HF-BN0-alg-e-05-0-UI.r1 NIH MGC 50 Homo sapiens cDNA clone
IMAGE:3079401 5', mRNA sequence.
ACCESSION
AW504318
VERSION
AW504318.1 GI:7141985
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 290)

REFERENCE
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab

CNDA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 32 Row: 1 Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
The following repetitive elements were found in this cDNA sequence:
8-63, >POLY A Simple repeat
Seq primer: M13 Forward.
FEATURES
source
Location/Qualifiers
1..290
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3079401"
/clone_lib="NIH_MGC_50"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="MGC85"
/lab_host="DH10B (LTI)"
/notes="Vector: pT73-Pac; Site 1: NotI; Site 2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 108 a 51 c 41 g 90 t
ORIGIN
Query Match 18.2%; Score 41; DB 10; Length 290;
Best Local Similarity 57.4%; Pred. No. 89;
Matches 74; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 31 GAAATAAATATGAGGCTCCATGATGCTATATGTTTCTTCTCAGAGAAATATTTA 90
DB 111 GAAAAAAGATATTTGGGCCCCAAAAATGGCATACGGTTTTATCCCAACCTAAGAAGAA 170
QY 91 GTTCTACTCAGGTTTTTCAAGCTAGCTGCTCCCAAAACGAAACAAACAAAAA 150
DB 171 GGATGTAATAAATAATTCGGCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 230
QY 151 CAACCTTTT 159
DB 231 CAAAATTTT 239

RESULT 9
BC022863
LOCUS
DEFINITION Homo sapiens, clone IMAGE:4077047, mRNA.
ACCESSION BC022863
VERSION BC022863.1 GI:221335650
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1562)
Strausberg, R.
Direct Submission
Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CNDA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amad@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

CNDA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 32 Row: 1 Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
The following repetitive elements were found in this cDNA sequence:
8-63, >POLY A Simple repeat
Seq primer: M13 Forward.
FEATURES
source
Location/Qualifiers
1..1562
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4077047"
/tissue_type="Bone marrow, acute myelogenous leukemia"
/clone_lib="NIH_MGC_55"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
BASE COUNT 478 a 289 c 331 g 464 t
ORIGIN
Query Match 18.2%; Score 41; DB 11; Length 1562;
Best Local Similarity 54.2%; Pred. No. 38;
Matches 83; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 1 TGATGTAGTGTCTTTCAGCATAAATTTGAAATAAATATATGAGGCTCCATATGCT 60
DB 1407 TGATAATATTTGAAAAATGGTGACCAAACTCTATGTAACCTTAGCAAAATTTTCATCT 1466
QY 61 ATATTGGTTTACCTTCAGAAGAAATATTAGTTTCTACTAGGTTTTCRAAGCTACGCTG 120
DB 1467 TTGTGCTTTGGAAATTTAGAAAATATTAAATTTCTTAAAGTCTTATAAAGTTTACTCA 1526
QY 121 TCCCCAAAAACGAAACCAACAAAAACAA 153
DB 1527 ATCCAAAAACAAAAACAAAAACAAAAA 1559

RESULT 10
AZ523166/c
LOCUS
DEFINITION 216PbB05 Pb MBN #21 Plasmodium berghei genomic 3', DNA sequence.
ACCESSION AZ523166
VERSION AZ523166.1 GI:13961890
KEYWORDS GSS.
SOURCE Plasmodium berghei.
ORGANISM Plasmodium berghei.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 595)
Carlton, J.M.-R. and Dame, J.B.
The Plasmodium vivax and P. berghei gene sequence tag projects
Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.
Location/Qualifiers
1..595
/organism="plasmodium berghei"
/strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
/db_xref="taxon:5821"
/clone_lib="Pb MBN #21"
/dev_stage="asexual blood forms"
/lab_host="Mus musculus"
/notes="Vector: pBluescript SK(+); vector DNA, phagemid excised from lambda ZAP; Site 1: EcoRV; Site 2: EcoRV; Genomic DNA was prepared from asynchronous blood stage forms of the cloned ANKA isolate of P. berghei grown in laboratory Swiss white mice. The DNA was purified from

contaminating host DNA by Hoechst Dye 33258-Cscl ultracentrifugation and precipitated. Purified DNA was digested with mung bean nuclease in the presence of 36-38% formamide at 50 C, as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6893-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size selected in the range 500-2000 bp. These were ligated into the EcoRV-cleaved and dephosphorylated pBluescript SK(+) vector. Recombinant plasmids were used to transform E. coli XL10-Gold host cells."

BASE COUNT 236 a 78 c 57 g 224 t
ORIGIN

Query Match 18.0%; Score 40.6; DB 17; Length 595;
Best Local Similarity 49.3%; Pred. No. 74;
Matches 106; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

11 TTGTTTCCAGGCATATAATTTGAATAAATATGAGGCTCCATGATATGCTATATGTTT 70
569 TTTTTCCTCCCAAAATTTAAATTTTAAATAATTAATAATTCCTCAATATGGAAAT 510

71 TACCTTCAGAGAATATTTAGTTTCTACTCAGGTTTTTCAAGCTACGCTGCCCAAAA 130
509 TTTTTCGAAAAAATAAATTTTTCCTGTTTAAAGAAAAAATTTTCCCTTAAAA 450

131 AACGAAACAAACAAAAACACCTTTTAAAGATTGATGGCTACTCATTTGATCTGCC 190
449 GGGGATCAAAAAAATAAATTTTTCCTGTTTAAAGAAAAAATTCCTCAAGAGGGGC 390

191 TCCTCTGCTGCAATCAATAGGAATTTTTTTTTTTT 225
389 ATATTTTTCCTGCAATTCATATGATATTAATTT 355

RESULT 11
CNS015W7/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN15B11 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
VERSION
AL105985.1 GI:5619413

KEYWORDS
SOURCE
Drosophila melanogaster.

ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 957)
Genoscope.

AUTHORS
TITLE
JOURNAL

COMMENT
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11

FEATURES
source
1. .957
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN15B11"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/notes="end : SP6"

274 a 139 c 139 g 318 t 87 others

BASE COUNT 274 a 139 c 139 g 318 t 87 others
ORIGIN

Query Match 18.0%; Score 40.6; DB 17; Length 595;
Best Local Similarity 49.3%; Pred. No. 74;
Matches 106; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Query Match 18.0%; Score 40.6; DB 17; Length 957;
Best Local Similarity 40.1%; Pred. No. 58;
Matches 73; Conservative 27; Mismatches 82; Indels 0; Gaps 0;

7 TAAAGTTGTTTCCAGGCATATAAATTTGAATAAATATGAGGCTCCATGATATGCTATATG 66
805 TWATTTTTTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 746

67 GTTTTACCTTCAGAGAATATTTAGTTTCTACTCAGGTTTTTCAAGCTACGCTGCCCCC 126
745 WTTTAAATGTTTCAMWAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGAAG 686

127 AAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 186
685 WATTAAATWAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 626

187 TG 188
625 KK 624

RESULT 12
AW381028/c
LOCUS
DEFINITION
CM1-HT0285-101199-036-h05 HT0285 Homo sapiens cDNA, mRNA sequence.

ACCESSION
AW381028
VERSION
AW381028.1 GI:6885687

KEYWORDS
SOURCE
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 525)
HCGP http://www.ludwig.org.br/ORESTES.

AUTHORS
TITLE
JOURNAL
COMMENT
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM1&t2=CM1-HT0285-101199-036-h05&t3=1999-11-10&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 36
High quality sequence stop: 112.
Location/Qualifiers
1. .525
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0285"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI; Site: 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 161 a 97 c 97 g 170 t
ORIGIN

Query Match 18.0%; Score 40.4; DB 10; Length 525;
Best Local Similarity 51.7%; Pred. No. 87;
Matches 92; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

FEATURES
source
1. .525
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0285"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI; Site: 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 161 a 97 c 97 g 170 t
ORIGIN

Query Match 18.0%; Score 40.4; DB 10; Length 525;
Best Local Similarity 51.7%; Pred. No. 87;
Matches 92; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

45 GGCTCCATGATATGCTATATGTTTACCTTCAGAGAATAATTTAGTTTCACTCAGGTT 104

Francisco
Box 0811, San Francisco, CA 94143-0811, USA
Tel: 415 206 8846
Fax: 415 206 3353
Email: malaris@itsa.ucsf.edu

Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.

Seq primer: T3

Class: shotgun

High quality sequence stop: 647.

Location/Qualifiers

FEATURES

source

1. .647
/organism="Cryptosporidium parvum"
/strain="IOWA"
/db_xref="taxon:5807"
/clone_lib="CpIOWAGDNa1"
/lab_host="E. coli XL2 Blue MRF"
/notes="Vector: pBlueScript II (SK-); Site 1: EcoRV; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center (http://sequence-www.stanford.edu/group/techdev/shear.htm)". The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoRV-digested, alkaline phosphatase-treated pBlueScript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF'. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

BASE COUNT 183 a 92 c 90 g 282 t

ORIGIN

Query Match 17.7%; Score 39.8; DB 17; Length 647;
Best Local Similarity 64.8%; Pred. No. 1e+02;
Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
Qy 81 AGAATATTAGTTTCACTCAGGTTTTCAAAGCTACGCTCTCCCAAAAACGAAACAA 140
Db 630 ATAGATTTCATTCACTAAAGTTTAGAAAGCTACTCTGCCTCAAAAAAAGCTCGAACAA 571
Qy 141 AACAAAAAACACCTTTTAAAGAGTTGATG 171
570 ATGAAACTATATGAAGATCAGATTCGATG 540

Search completed: June 9, 2003, 10:23:43
Job time : 1060 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic : nucleic search, using sw model

Run on: June 7, 2003, 14:33:54 ; Search time 154 Seconds
(without alignments)
3290.257 Million cell updates/sec

Title: US-09-765-231a-58

Perfect score: 225

Sequence: 1 tgaatgtagtgcgttcagg.....attagaattttttttttttt 225

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	225	100.0	225	AAH23128	Osteoarthritis tis
2	48.4	21.5	1664976	AAV21209	Methanococcus jann
3	41	18.2	5518	ABK28306	DNA transcription
4	40	17.8	6809	ABN80174	Human chemically m
5	40	17.8	6809	ABL70557	Chemically treated
6	40	17.8	6809	AA661214	Human gene regulat
7	40	17.8	6809	ABK31314	Signal transductio
8	38.2	17.0	640681	ABK922787	Buchnera sp. genom
9	38	16.9	12393	ABL33263	Human immune syste

C 10	37.4	16.6	2270	20	AAH23128	Human protein phos
C 11	37.2	16.5	778	22	ABL21862	Human breast cance
C 12	36.6	16.3	2705	19	AAV05164	Human growth arres
C 13	36.4	16.2	345	24	ABQ85902	Arabidopsis thalia
C 14	36.4	16.2	1479	24	ABN95163	Gene #1661 used to
C 15	36.4	16.2	7924	24	ABK40070	Human chemically p
C 16	36.4	16.2	7924	24	ABL34139	Human immune syste
C 17	36.4	16.2	40388	23	ABL26342	Drosophila melanog
C 18	36.2	16.1	5474	24	ABL33270	Human immune syste
C 19	36	16.0	5153	17	AAH230347	Human YAP CDNA. H
C 20	36	16.0	6725	24	ABL33208	Human immune syste
C 21	36	16.0	6725	24	ABL34554	Human metastasis a
C 22	35.8	15.9	502	24	ABO88918	Human prostate exp
C 23	35.8	15.9	3002	23	ABL15414	Drosophila melanog
C 24	35.8	15.9	7847	24	ABL34188	Human immune syste
C 25	35.6	15.8	6163	24	ABN80119	Human chemically m
C 26	35.4	15.7	932	22	AAK88641	Human digestive sy
C 27	35.4	15.7	113515	24	ABL34175	Human immune syste
C 28	35.2	15.6	557	22	ABL18969	Human breast cance
C 29	35.2	15.6	11049	24	ABL92218	Chemically treated
C 30	35.2	15.6	11049	24	ABL49321	Human polynucleoti
C 31	35.2	15.6	11049	24	ABL32668	Human immune syste
C 32	35	15.6	332	22	AAK55869	Human immune/haema
C 33	35	15.6	1132	23	ABV22830	Human prostate exp
C 34	35	15.6	1132	23	ABV28660	Human prostate exp
C 35	35	15.6	3541	19	AAV07076	CDNA encoding huma
C 36	35	15.6	3541	24	AAH38589	Human stromal cell
C 37	35	15.6	10609	22	AAH46359	Tumour suppressor
C 38	35	15.6	10609	24	ABK31268	Signal transductio
C 39	34.8	15.5	355	22	AAH81841	Rat differential t
C 40	34.8	15.5	388	22	AAH89390	Human polynucleoti
C 41	34.8	15.5	5941	24	ABL33350	Human immune syste
C 42	34.8	15.5	6062	24	ABL34078	Human immune syste
C 43	34.8	15.5	910715	20	AAH20248	Borrelia burgdorfe
C 44	34.6	15.4	5216	24	ABL32347	Human immune syste
C 45	34.6	15.4	5216	24	ABL34459	Human metastasis a

ALIGNMENTS

RESULT 1
AAH23128
ID AAH23128 standard; DNA; 225 BP.
XX
AC AAH23128;
XX
DT 17-SEP-2001 (first entry)
XX
DE Osteoarthritis tissue-derived nucleic acid sequence #58.
XX
KW Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;
KW wound healing; osteopathic; anti-arthritis; anti-inflammatory;
KW vulnarary; antibacterial; antiallergic; ds.
XX
OS Homo sapiens.
XX
PN WO200153531-A2.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US00016.
XX
PR 18-JAN-2000; 2000US-0176523.
XX
PA (PHAA) PHARMACIA CORP.
XX
PI Phippard D, Vasanthakamur G, Dotson S, Ma X;
XX
DR WPI; 2001-451914/48.
XX
PT Substantially purified protein, polypeptide or their fragments, used to identify a biologically active compound or composition and treat

PT mammalian osteoarthritis -
XX
PS Claim 1; Page 137; 144pp; English.
XX
CC Sequences AH23071-23152 represent nucleic acid sequences derived from
CC osteoarthritis tissues. The sequences are useful as probes and for the
CC diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides
CC and polypeptides of the invention are useful for generating diagnostic
CC reagents, as targets for small molecule drug development, generation of
CC therapeutics, and cloning genes. Specific antibodies are used to generate
CC enzyme linked immunosorbent assays for detection of osteoarthritis. The
CC invented molecules can be used to treat osteoarthritis or to analyse the
CC disease-modifying activity of osteoarthritis drugs. Other disorders
CC treatable using the nucleic acid sequences include atopic, inflammatory
CC and infectious disorders e.g. Crohn's disease and sepsis, and wound
CC healing.
XX
SQ Sequence 225 BP; 72 A; 38 C; 35 G; 80 T; 0 other;
Query Match 100.0%; Score 225; DB 22; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.2e-45;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGATGTTAGTCTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60
DB 1 TGATGTTAGTCTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60
QY 61 ATATTGTTTACCTTCAGAGATATTTAGTTTCACTCAGGTTTTCAGAGCTAGCTG 120
DB 61 ATATTGTTTACCTTCAGAGATATTTAGTTTCACTCAGGTTTTCAGAGCTAGCTG 120
QY 121 TCCCCAAAAACGAAACAAAAACAAAAACAACTTTTAAAGATTGATGGCTACTCAT 180
DB 121 TCCCCAAAAACGAAACAAAAACAAAAACAACTTTTAAAGATTGATGGCTACTCAT 180
QY 181 TTGATGCTCCTCTCTGCTGAATCAATAGGAATTTTTTTTTT 225
DB 181 TTGATGCTCCTCTCTGCTGAATCAATAGGAATTTTTTTTTT 225
RESULT 2
AAV21209
ID AAV21209 standard; DNA; 1664976 BP.
XX
AC AAV21209;
XX
DT 10-NOV-1998 (first entry)
Methanococcus jannaschii circular chromosome.
KW Methanococcus jannaschii; methanogenic archaeon; circular chromosome;
KW genome; autotrophic; extrachromosomal element; identification; ds.
XX
OS Methanococcus jannaschii.
XX
XX WO9807830-A2.
XX
XX 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-US14900.
XX
XX 22-AUG-1996; 96US-0024428.
XX
XX (GENO-) INST GENOMIC RES.
PA (UNII) UNIV ILLINOIS FOUND.
PA (UJYO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX Bult CJ, Smith HO, Venter JC, White OR, Woese CR;
XX
XX WPI; 1998-169145/15.
XX
XX Complete genome sequence of methano-genic archaeon, Methanococcus
PT jannaschii - useful in identification of M. jannaschii genome

PT fragment
XX
PS Claim 13; Page 152-585; 614pp; English.
XX
CC The present sequence represents the complete 1.66-megabase pair genome
CC sequence of the Methanococcus jannaschii circular chromosome. The
CC present invention describes M. jannaschii open reading frames from the
CC genome sequence. The invention also describes a computer based system
CC for identifying fragments of the M. jannaschii genome that are
CC homologous to target nucleotide sequences, comprising: (a) data storage
CC means comprising the nucleotide sequence of the 1664976, 58407 or 16550
CC bp sequence (see AAV21209, AAV21210 and AAV21211), or a nucleotide
CC sequence at least 99.9% identical to it; (b) search means for comparing a
CC target sequence to the nucleotide sequence of the data storage means to
CC identify a homologous sequence, and (c) retrieval means for obtaining
CC the homologous sequence. The method, which is based on whole genome
CC random sequencing of an autotrophic archaeon M. jannaschii, the genome
CC of which consists of 3 physically distinct elements, a large circular
CC chromosome (the 1664976 bp sequence given in AAV21209), a large circular
CC extra-chromosomal element (the 58407 bp sequence given in AAV21210), and
CC a small circular extra-chromosomal element (the 16550 bp sequence given
CC in AAV21211), can be used in the identification of M. jannaschii genome
CC fragment.
XX
SQ Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T; 101 other;
Query Match 21.5%; Score 48.4; DB 19; Length 1664976;
Best Local Similarity 51.9%; Pred. No. 0.029;
Matches 109; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 7 TAAGTTGTTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCTATATG 66
DB 151403 TAAATCTCTGAGCCCTCTTATTTCCATCAATATTATAGTACATGTTTGTGTAATTTG 151462
QY 67 GTTTTACCTTCAGAGATATTTAGTTTCACTCAGGTTTTCAGAGCTAGCTGTCCTCC 126
DB 151463 TCTTTATCAATAGCCCTTCTTAAATCTCTCTGTAATTTTAAAACTCCGAGTTGTCT 151522
QY 127 AAAAAACGAAACAAAAACAAAAACAACTTTTAAAGATTGATGGCTACTCATTTGATC 186
DB 151523 AAAAACTATAAATAAATCAATATTACTCTTTTCCATACTCTTTAATTTGTTTATC 151582
QY 187 TGCTCTCTCTGCTGAATCAATAGGAATTT 216
DB 151583 TTACCCAAAAATCCCACTTATTAGGAATTT 151612
RESULT 3
ABK28306/c
ID ABK28306 standard; DNA; 5518 BP.
XX
AC ABK28306;
XX
DT 23-APR-2002 (first entry)
XX
DE DNA transcription associated complementary genomic DNA #90.
XX
XX DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
KW immunological disorder; Werner syndrome; developmental disorder;
KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
KW polyglutamine disorder; solid tumour.
XX
XX Unidentified.
XX
XX WO200192565-A2.
XX
XX 06-DEC-2001.

XX 06-APR-2001; 2001WO-EP03973.
PF XX
XX 06-APR-2000; 2000DE-1019058.
PR XX
PR 07-APR-2000; 2000DE-1019173.
PR XX
PR 30-JUN-2000; 2000DE-1032529.
PR XX
PR 01-SEP-2000; 2000DE-1043826.
XX XX
XX (EPIG-) EPIGENOMICS AG.
PA XX
PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-090046/12.
DR XX
XX New nucleic acids or oligomers, useful for diagnosing or treating
PT diseases associated with DNA transcription, e.g. immunological
PT disorders, Werner syndrome, psoriasis, myocardial infarction, solid
PT tumours or cancer
XX
XX Claim 1; SEQ ID No 180; 32pp; English.
XX
XX The invention relates to a nucleic acid, which comprises a segment of the
CC chemically pretreated DNA of genes associated with DNA transcription from
CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
CC to the chemically pretreated DNA of genes associated with DNA
CC transcription. The set of oligomer probes are useful for detecting the
CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
CC diagnosing or treating diseases associated with DNA transcription
CC (particularly with the methylation status), e.g. adenosine deaminase
CC deficiency, viral infection, retroviral infection, Sezary syndrome,
CC haematological disorders, immunological disorders, Werner syndrome,
CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
CC neurological disorders, neurodegenerative disorders, Waardenburg
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
CC associated genomic DNA molecules of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 5518 BP; 1899 A; 52 C; 1031 G; 2536 T; 0 other;
Query Match 18.2%; Score 41; DB 24; Length 5518;
Best Local Similarity 51.4%; Pred. No. 0.87; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 7 TAAAGTTGTTTCAGGCATAAATTTGAAATAAATTTATGAGGCTCCATGATATGCTATATTG 66
Db 4088 TATATTTTAAACAAAAATTTTCATTTAACTATAAATCTCTTAAATAATTAACGAA 4029
QY 67 GTTTTACCTTCAGAGAAATTTAGTTTCTACTCAGGTTTTCAGAGCTAGCTGTCCTCCC 126
Db 4028 TTACTTAAATTAATAACTAAATAATAATACTTCAAACTCTTAAAACTAATTAATCTCTAA 3969.
QY 127 AAAAAACGAACAAACAAACAAACCTTTTAAAGTTGATGGCTACTCTTTGATC 186
Db 3968 AATAAAATACTTAATAATAAAACACTATTTTAAATTAATAACAACTTAATCTTATT 3909
QY 187 TGCCT 191
Db 3908 TAACT 3904
RESULT 4
ABN80174/c
ID ABN80174 standard; DNA; 6809 BP.
XX
AC ABN80174;
XX

DT 15-JUL-2002 (first entry)
XX Human chemically modified disease associated gene SEQ ID NO 191.
DE XX
XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytostatic; anticonvulsant; ds.
XX
XX Homo sapiens.
OS Synthetic.
OS
PN WO200200927-A2.
XX
XX 03-JAN-2002.
PD
XX 02-JUL-2001; 2001WO-EP07536.
PF
XX 30-JUN-2000; 2000DE-1032529.
PR
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
PA
PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130908/17.
DR
XX Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes, comprises a
PT sequence of a segment of chemically pretreated DNA of genes associated
PT with development
XX
PS Claim 1; SEQ ID NO 191; 27pp; English.
XX
XX The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in
CC the specification such as ACCPN, ADFN, or AFD1 and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Currarino syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs).
CC Note: The sequence data for this patent did not form part of the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 6809 BP; 1792 A; 270 C; 1625 G; 3122 T; 0 other;
Query Match 17.8%; Score 40; DB 24; Length 6809;
Best Local Similarity 54.9%; Pred. No. 1.6;
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 17 CAGGCATAAAATTTGAAATAAATTTATGAGGCTCCATGATATGCTATATTGTTTACCTT 76
Db 1384 CACANATAAAACCAATATATATCGTAACACACACTCTATTTCATAATTTCTTAATTAT 1325
QY 77 CAGAAGAATATTAGTTTCTACTCAGGTTTTCAGAGCTAGCTGTCCTCCCAAAACGAA 136
Db 1324 ACAAAACACTATTAAATTAATCCCGTCTATCAATTTTCGTTTACTACAAATTACTTT 1265
QY 137 ACAAACAAACAAACAAACCTTTT 160
Db 1264 TCAAACTTAATCAATAAATCTTT 1241

RESULT 5
ABL70557/c
ID ABL70557 standard; DNA; 6809 BP.
XX AC ABL70557;
XX DT 01-JUL-2002 (first entry)
XX DE Chemically treated cell signalling DNA sequence#224.
XX KW Cell signalling; cytosine methylation; cell signalling disease;
KW cancer; tumour; cytostatic; ds.
XX OS Unidentified.
XX PN WO200202807-A2.
XX DT 10-JAN-2002.
XX PR 29-JUN-2001; 2001WO-EP07471.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-154758/20.
XX PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signalling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signalling
XX PS Claim 1; SEQ ID NO 447; 24pp-sequence listing; English.
XX CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX SQ Sequence 6809 BP; 1792 A; 270 C; 1625 G; 3122 T; 0 other;
Query Match 17.8%; Score 40; DB 24; Length 6809;
Best Local Similarity 54.9%; Pred. No. 1.6;
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 17 CAGGCATATAAATTGAAATTAATATGAGGCTCCATGATGCTATATGTTTACCTT 76
DB 1384 CACAATAAACCACATATATATCGTAACACACACTCTATTCTTAATTAT 1325
QY 77 CAGAAGATATTTAGTTTCTACTCAGGTTTTTCAAAGCTACGCTGTCGCCCAAAAACGAA 136
DB 1324 ACAAAACTATTTAAATTAATCCGCTCTATCAATTTTCGTTTACTACAAATTACTTT 1265
QY 137 ACAAAACAAAACACCTTTT 160
DB 1264 TCAAACTTAATCATAAATCTTT 1241

RESULT 6
AAS61214/c
ID AAS61214 standard; DNA; 6809 BP.
XX AC AAS61214;
XX DT 29-JAN-2002 (first entry)
XX DE Human gene regulation-associated gene oligonucleotide #169.
XX KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
KW cardiac damage; inflammatory response; haemophilia; Werner syndrome;
KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
KW renal disease; Preclampsia; cardiac allograft vascular disease;
KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
XX OS Homo sapiens.
XX PN WO200177375-A2.
XX PD 18-OCT-2001.
XX PF 06-APR-2001; 2001WO-EP03968.
XX PR 06-APR-2000; 2000DE-1019058.
XX PR 07-APR-2000; 2000DE-1019173.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-017470/02.
XX PT New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease
XX PS Claim 1; SEQ ID No 173; 26pp; English.
XX CC The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 6809 BP; 1792 A; 270 C; 1625 G; 3122 T; 0 other;
Query Match 17.8%; Score 40; DB 24; Length 6809;
Best Local Similarity 54.9%; Pred. No. 1.6;
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 17 CAGGCATATAAATTGAAATTAATATGAGGCTCCATGATGCTATATGTTTACCTT 76

Db 198302 AATCAAAATTAATAAGTAGTATATTTTCAATTATGACTTTTCAATAATTTTATTATTAAAG 198243
 QY 92 TTTCAGTCAAGTTTTCACAGCTAGCTGTGCCCAAAACGAAACAAAACAAAAC 151
 Db 198242 TAATAATTTGTTATTTATAATAACTTTTCACTAAAATAAACAATACAAAACAAAACAAAAC 198183
 QY 152 AACCTTTTAA 162
 Db 198182 AAGATTTTAA 198172

RESULT 9
 ABL33263/c
 ID ABL33263 standard; DNA; 12393 BP.
 XX ABL33263;
 AC ABL33263;
 XX 26-MAR-2002 (first entry)
 DT Human immune system associated gene SEQ ID NO: 1236.
 XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianemic; cytosine; cytosine; cytosine;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;
 KW acute myeloid leukemia; cancer; eye disease; arteriosclerosis; anaemia;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

XX Homo sapiens.
 OS Homo sapiens.
 XX WO200200928-A2.
 PN 03-JAN-2002.
 PD 02-JUL-2001; 2001WO-EP07537.
 PF 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 DR Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 XX cytosine methylation
 PS Claim 1; SEQ ID NO 1236; 32pp + Sequence Listing; German.
 XX

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 CC
 XX Sequence 12393 BP; 3484 A; 219 C; 2406 G; 6282 T; 2 other;
 SQ

Query Match 16.9%; Score 38; DB 24; Length 12393;
 Best Local Similarity 49.5%; Pred. No. 5.2;
 Matches 98; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 QY 22 ATAAATTTGAATAATTATGAGCTCCATGATATCTATATGTTTACCTTCAGAA 81
 Db 6387 AAAAACTTAAAAAAAATTTTAAAAAATAATTTAAAAAATTTTCTCTTAAAA 6328

QY 82 GAATATTTAGTTTCACTCAGGTTTTTCAAGCTACGCTGTCCCAAAACGAAACAAA 141
 Db 6327 AAAACATAATTTTATTTAACGATTTTAAAAAATCAAAATACTCGATAAATAATAA 6268
 QY 142 AAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATCTGCTCTCTGCTGA 201
 Db 6267 TAAATAAAAAATCAATATACTCAATTTCTCCCGCTCTCTCAATAAACCTTA 6208
 QY 202 ATCAATTAGGAATTTT 219
 Db 6207 AACAAATTTCTATCATTT 6190

RESULT 10
 AAX05715/c
 ID AAX05715 standard; DNA; 2270 BP.
 XX AAX05715;
 AC AAX05715;
 XX 07-MAY-1999 (first entry)
 DT Human protein phosphatase (PROPHO) encoding DNA.
 XX Human protein phosphatase (PROPHO) encoding DNA.
 KW Protein phosphatase; PROPHO; apoptosis; AIDS; Alzheimer's Disease;
 KW Acquired Immune Deficiency Syndrome; Parkinson's Disease; inflammation;
 KW cell proliferation; Addison's disease; allergy; anemia; cancer; bone;
 KW leukemia; breast; brain; human; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key
 FH Location/Qualifiers
 FT 70..1509
 CDS /*tag= a
 FT /product= "protein phosphatase (PROPHO)"
 FT
 XX

PN WO9856925-A1.
 XX 17-DEC-1998.
 PD 11-JUN-1998; 98WO-US11614.
 PF 11-JUN-1997; 97US-0873093.
 PR (INCY-) INCYTE PHARM INC.
 XX Bandman O, Corley NC, Goli SK, Lal P, Zhang H;
 XX WPI; 1999-080906/07.
 DR P-PSDB; AAW94283.
 XX New substantially purified human protein phosphatase (PROPHO) -
 PT useful in the diagnosis, prevention or treatment of inflammation,
 PT cancer, and disorders associated with apoptosis
 XX Claim 5; Fig 1A-G; 73pp; English.

PS This DNA encodes a human protein phosphatase (PROPHO). Host cells
 XX containing a vector comprising the PROPHO nucleic acid are used for the
 CC recombinant production of the protein. PROPHO forms a composition in the
 CC treatment or prevention of apoptosis-related disorders (e.g. Acquired
 CC Immune Deficiency Syndrome (AIDS), Alzheimer's Disease and Parkinson's
 CC disease), and in the stimulation of cell proliferation. Antagonists of
 CC the protein are useful in treating inflammation (e.g. Addison's disease,
 CC allergies and anemia), and disorders associated with cell proliferation
 CC (including various cancers like leukemia, and cancers affecting bone,
 CC breast and brain). Complementary polynucleotides are useful in detecting
 CC polynucleotides that encode PROPHO, useful in the diagnosis of conditions
 CC associated with the expression of PROPHO, and in assays that detect
 CC activation or induction of various cancers. PROPHO is useful in producing
 CC antibodies or screening libraries of pharmaceutical agents in order to
 CC identify those that bind to PROPHO.
 XX Sequence 2270 BP; 693 A; 344 C; 515 G; 716 T; 2 other;
 SQ

Query Match 16.6%; Score 37.4; DB 20; Length 2270;
Best Local Similarity 53.0%; Pred. No. 5.8;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 29 TTGAATAAATATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGAGAATATT 88
DB 2197 TTGCATACATTTTGATGTGTGTAGTTAGGCTATACAGATATACAAACCAAGCTGTAT 2138

QY 89 TAGTTTCACCTCAGGTTTTCAAGCTACGCTGTCCCAAAACGAAACAAACAAAA 148
DB 2137 GAATAATACATAGGTTTTCAAAAGTTATGTTTCATAAAGAATACAGAGAAGGAAAAAT 2078

QY 149 AACCAACCTTTTAAGAGTTGATGCTACTCA 179
DB 2077 AGCCACACATCCAATATCTCACAACTTCTAA 2047

RESULT 11

AA121862/C
AA121862 standard; cDNA; 778 BP.

AA121862;
07-DEC-2001 (first entry)
Human breast cancer expressed polynucleotide 14319.
Human; breast cancer; cell marker; cytostatic; ss.
Homo sapiens.

WO200151628-A2.

19-JUL-2001.

10-JAN-2001; 2001WO-US00798.

14-JAN-2000; 2000US-0176077.

14-MAR-2000; 2000US-0189167.

24-MAR-2000; 2000US-0192099.

29-MAR-2000; 2000US-0193480.

15-MAY-2000; 2000US-0205230.

09-JUN-2000; 2000US-0211315.

25-JUL-2000; 2000US-0220534.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Lillie J, Xu Y, Wang Y, Steinmann K;

WPI; 2001-451856/48.

New peptide useful as a marker for the diagnosis of breast cancer -

Claim 1; Page 2564-2565; 3695pp; English.

The invention relates to human breast cancer expressed polynucleotides (AA107544-AA126789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity.

Sequence 778 BP; 220 A; 134 C; 180 G; 244 T; 0 other;

Query Match 16.5%; Score 37.2; DB 22; Length 778;
Best Local Similarity 55.4%; Pred. No. 5.7;
Matches 72; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 27 ATTTGAATAAATATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGAAGATA 86

Db 136 AATTTAGCTATATTATCTCTCAATCAAAATTTCTTCTCAGAGTAAGTAAAAAGAGT 77
QY 87 TTTAGTTTCACTCAGGTTTTTCAAAGCTAGCTGTCCCAAAACGAAACAAACAAA 146
Db 76 TATAGCAATACATATAATGGACAAAAGGTATCCCAAAAAAAGAAAAA 17
QY 147 AAAACAACCT 156
Db 16 AAAAGTACCT 7

RESULT 12

AAV05164
ID AAV05164 standard; cDNA; 2705 BP.

AAV05164;

06-JUL-1998 (first entry)

Human growth arrest gene B4B.

Growth arrest gene; B4B gene; cell proliferation; cancer; leukaemia;
marker; gene therapy; human; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 123..596

/*tag= a

/note= "(Claim 4)"

WO9800544-A2.

08-JAN-1998.

27-JUN-1997; 97WO-US11341.

28-JUN-1996; 96US-0021874.

(ACTI-) ACTIVATED CELL THERAPY INC.

Engleman EG, Laus R, Ruegg CL;

WPI; 1998-086969/08.

P-PSDB; AAW46489.

New isolated DNA encoding growth arresting protein expressed in B cells - and related vectors, transformed cells and proteins, used for, e.g. gene therapy of leukaemia or diagnosis of cancer and immune status

Claim 4; Fig 1; 26pp; English.

This cDNA clone, the coding region of which is claimed, codes for a growth arrest protein (see AAW46489). It corresponds to the B4B gene, expression of which results in inhibition of cellular proliferation. Human peripheral blood mononuclear cells were fractionated to recover high, intermediate and low density fractions. cDNA libraries from all 3 fractions were subjected to a differential display PCR. Bands unique to the intermediate density sample were reamplified, cloned and sequenced. A 101 bp insert was used to screen a cDNA library from intermediate density cells to isolate the 2.7 kb cDNA clone. This insert was used to isolate a genomic fragment localised to the B4B gene to chromosome 20q12-q13.1. The isolated clone can be used to express the growth arrest protein in host cells. Introduction of a functional B4B allele (by gene therapy) into a cell with a defective allele may be used to prevent development of pro-B-cell leukaemia or other malignancies. Introduction of the isolated DNA in a vector that includes a suicide inhibition cassette can be used to arrest expansion of cells, in vivo or in vitro, e.g. useful in skin cell grafts and in lymphocytes containing a heterologous

CC cytokine gene for in vivo expression.

XX Sequence 2705 BP; 729 A; 584 C; 627 G; 765 T; 0 other;

Query Match 16.3%; Score 36.6; DB 19; Length 2705;

Best Local Similarity 57.4%; Pred. No. 9.3;

Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY. 39 TTATGAGGCTCCATGATATCTATATTGGTTTACCTTCAGAGAATATTAGTTTCACT 98

DB 2588 TTAATAGATTATTTCATATCTATATTGTAATATTTCATACAAATGTTTATACTCT 2647

QY 99 CAGGTTTTTCAAGCTACGCTGTCCTCCCAAAACGAAACAAACAAAAACAA 153

DB 2648 AGGATATAAAACAGATTCTGATTCCTCCCAAAACGAAACAAACAAAAACAA 2702

RESULT 13

ABQ85902/c

ABQ85902 standard; DNA; 345 BP.

ABQ85902;

05-SEP-2002 (first entry)

Arabidopsis thaliana expressed polynucleotide SEQ ID NO 772.

Plant; Arabidopsis; transgenic; fungicide; insecticide; ds.

Arabidopsis thaliana.

US2002062014-A1.

23-MAY-2002.

26-JAN-2001; 2001US-0770791.

27-JAN-2000; 2000US-178480P.

(GOL/) GORLACH J.

(ANY/) AN Y.

(HAMI/) HAMILTON C M.

(PRIC/) PRICE J L.

(RAIN/) RAINES T M.

(YUY/) YU Y.

(RAME/) RAMEKA J G.

(PAGE/) PAGE A.

(MATH/) MATHW A V.

(LEDF/) LEDFORD B L.

(WOES/) WOESSNER J P.

(HAAS/) HAAS W D.

(GARC/) GARCIA C A.

(KRIC/) KRICKER M.

(SLAT/) SLATER T.

(DAVI/) DAVIS K R.

(ALLE/) ALLEN K.

(HOFE/) HOFFMAN N.

(HURB/) HUREAN P.

Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;

Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;

Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;

Hurban P;

WPI; 2002-479265/51.

New nucleic acid sequences of Arabidopsis thaliana and their encoded products are useful to produce transgenic plants, to screen for biologically active agents such as fungicides and insecticides and in genetic studies

Claim 1; SEQ ID NO 772; 18pp + Sequence Listing; English.

CC The invention relates to a novel nucleic acid of Arabidopsis thaliana comprising a sequence capable of hybridizing under stringency to one of the 999 sequences referred to but not defined in the specification (AB08131-ABQ86129). The nucleic acid sequences are useful to identify CC homologous or related genes, to produce compositions that modulate CC expression or function of the encoded protein, to map functional regions CC of the protein, to study associated physiological pathways, to CC genetically manipulate cells and plants. The encoded products are useful CC to screen for biologically active agents such as fungicides or CC insecticides and to elucidate biochemical pathways.

XX Sequence 345 BP; 97 A; 61 C; 48 G; 139 T; 0 other;

QY Query Match 16.2%; Score 36.4; DB 24; Length 345;

DB Best Local Similarity 53.5%; Pred. No. 8;

Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 12 TGTTCAGGCATAAATTTGAAATAATATATGAGGCTCCATGATATGCTATTTGTTTT 71

DB 155 TGTATTACCTAAGACAATAATATGATTTATGATAAAAACTAATTTGAAATTACTGCATAT 96

QY 72 ACCTTCAGAGAATATTTAGTTTCACTCAGGTTTTTCAAGCTACGCTGCTCCCAAAAA 131

DB 95 GAAATATATAATCTTTTGTGTTTAATCTCAAGATGTAATAATTATGATCTTCACATAAA 36

QY 132 ACGAAACAAACAAAAACAA 153

DB 35 AAAAAAAAAAAAAAAAAAAAAA 14

RESULT 14

ABN95163

ID ABN95163 standard; DNA; 1479 BP.

AC ABN95163;

DT 13-AUG-2002 (first entry)

XX Gene #1661 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

KW metastatic liver tumour; cytostatic; expression profile; disease state;

KW disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

XX WO200229103-A2.

PN 11-APR-2002.

PD 02-OCT-2001; 2001WO-US30589.

PF 02-OCT-2000; 2000US-237054P.

PR (GENE-) GENE LOGIC INC.

PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

DR Diagnosing and detecting the progression of liver cancer,

XX hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample

XX Claim 1; SEQ ID NO 1661; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the CC progression of liver cancer, hepatocellular carcinoma or metastatic liver CC tumor in a patient, and differentiating metastatic liver cancer from CC hepatocellular carcinoma in a patient, involving detecting the level of CC expression of two or more genes represented in ABN93503-ABN97455 in a CC tissue sample. The method of the invention has hepatotropic, and

CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer. Hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1479 BP; 432 A; 287 C; 266 G; 493 T; 1 other;

Query Match 16.2%; Score 36.4; DB 24; Length 1479;
Best Local Similarity 61.7%; Pred. No. 9.6;
Matches 58; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
OY 128 AAAAAAGAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATGCTACTCTATTGATCT 187
DB 834 AGAAGAAGAAACAAACAAACAAACAAACCTTTTAAATATTACTTGCACCTCAACAGACT 893
DB 188 GCCTCCTCTGCTGAATCAATTAGGAATTTTTTTT 221
894 CCCTGCGTACTGCTCTTTCCAGGAACCTTTACTT 927

RESULT 15
ABK40070/c
ID ABK40070 standard; DNA; 7924 BP.
XX
AC ABK40070;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #76 strand 2.
XX
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW Cytostatic; ALDH6; CYP11A; CYP11B; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200202806-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07470.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.

(EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154757/20.
XX
XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT useful for detecting cytosine methylation state of genes associated
PT with pharmacogenomics and for therapy of diseases e.g. cancer
XX
PS Claim 1; SEQ ID No 152; 24pp; English.

XX The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated DNA
CC of genes associated with pharmacogenomics according to one of the
CC sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC MRP (NM_004996, NM_015900, NM_019901, NM_019902, NM_019862, NM_019898,
CC NM_019859) and their complementary sequences, or a sequence (S1) chosen
CC from 87 sequences and their complementary sequences. The chemical pretreatment
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)
CC into uracils. Also included are an oligomer (II) in particular an

CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC each case at least one base sequence having a length of 9 nucleotides
CC which hybridises to or is identical to a chemically pretreated DNA of
CC genes associated with pharmacogenomics and their complements, arranged in
CC an array for analysing diseases associated with the methylation state
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 7924 BP; 1973 A; 83 C; 1827 G; 4041 T; 0 other;
Query Match 16.2%; Score 36.4; DB 24; Length 7924;
Best Local Similarity 64.0%; Pred. No. 12;
Matches 55; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
OY 77 CAGAAAGAAATATTAGTTTCACTCAGGTTTTTCAAGCTACGCTGCCCCCAAAACGAA 136
DB 3179 CAATTAATTTTACCTCTACTCATATATATCAATTAACCTTCCCAATTAACAACA 3120
OY 137 ACAAAACAAACAAACAAACCTTTTAA 162
DB 3119 ATAAACGAAACAAACGACCTTATTCA 3094

Search completed: June 7, 2003, 20:37:22
Job time : 299 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2003, 17:27:51 ; Search time 898 Seconds
(without alignments)
7291.906 Million cell updates/sec

Title: US-09-765-231A-58
Perfect score: 225
Sequence: 1 tcatggtagttgtttcagg.....attaggaattttttttttt 225

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sta.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_red.*
- 36: em_htg_rnd.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	225	100.0	225	6	AX202128	Sequence
C 2	225	100.0	143800	2	AC011875	Homo sapi
C 3	225	100.0	186962	9	AP000848	Homo sapi
C 4	175.6	78.0	172830	2	AP001320	Homo sapi
C 5	128.4	57.1	75002	2	AC023384	Homo sapi
C 6	47.2	21.0	30726	2	AC117269	Dictyoste
C 7	46.2	20.5	192454	2	AC094553	Rattus no
C 8	45.2	20.1	9971	1	U67577	Methanococ
C 9	44	19.6	146353	2	AC104893	Mus muscu
C 10	43.8	19.5	200110	10	AC117237	Mus muscu
C 11	42.6	18.9	128526	2	AP005406	Oryza sat
C 12	42.4	18.8	199551	2	AC006281	Plasmodiu
C 13	42	18.7	147760	9	AC011846	Homo sapi
C 14	41.8	18.6	156140	9	AC067745	Homo sapi
C 15	41.8	18.6	215467	2	AC013420	Homo sapi
C 16	41.8	18.6	215734	2	AC073828	Mus muscu
C 17	41.6	18.5	171050	2	AC112232	Homo sapi
C 18	41.6	18.5	181864	2	AC027460	Homo sapi
C 19	41.2	18.3	235776	2	AL645950	Mus muscu
C 20	41	18.2	5518	6	AX323692	Sequence
C 21	41	18.2	169546	2	AC004157	Plasmodiu
C 22	40.8	18.1	62354	2	AC118028	Mus muscu
C 23	40.8	18.1	130355	8	AP003412	Oryza sat
C 24	40.6	18.0	203241	9	AC008282	Homo sapi
C 25	40.4	18.0	12300	10	MMU84903	Mus musculu
C 26	40.4	18.0	101491	10	AP003183	Mus muscu
C 27	40.4	18.0	107739	2	AC116979	Dictyoste
C 28	40.4	18.0	192107	2	AC107635	Mus muscu
C 29	40.4	18.0	201773	2	AC130714	Mus muscu
C 30	40.4	18.0	210614	2	AC013548	Mus muscu
C 31	40.4	18.0	252504	2	AC073815	Mus muscu
C 32	40.2	17.9	122032	9	AL513187	Human DNA
C 33	40.2	17.9	161067	2	AL607133	Homo sapi
C 34	40.2	17.9	208398	2	AC123245	Rattus no
C 35	40.2	17.9	213728	9	AL158828	Human DNA
C 36	40	17.8	6809	6	AX251912	Sequence
C 37	40	17.8	6809	6	AX344310	Sequence
C 38	40	17.8	6809	6	AX344766	Sequence
C 39	40	17.8	6809	6	AX348989	Sequence
C 40	40	17.8	67970	3	PFMAL1P3	Plasmodiu
C 41	40	17.8	68751	2	AC103861	Homo sapi
C 42	40	17.8	153826	2	AC083965	Homo sapi
C 43	40	17.8	163608	2	AC026081	Homo sapi
C 44	40	17.8	166379	9	AC087378	Homo sapi
C 45	40	17.8	176975	2	AC098753	Rattus no

ALIGNMENTS

RESULT 1
AX202128
LOCUS AX202128 225 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 58 from Patent WO0153531.
ACCESSION AX202128
VERSION AX202128.1 GI:15391919
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 225)
AUTHORS Phippard,D., Vasanthakumari,G., Doteon,S. and Ma,X.J.
TITLE Osteoarthritis tissue derived nucleic acids, polypeptides, vectors, and cells

2

Db 50974 TGATGGTAAGTTGTTTCAGGCATAAAATTGAAATAAAATTATGAGGCTCCATGATATGCT 50915

```

AP001320      172830 bp      DNA      linear      HTG 30-MAY-2000
LOCUS      Homo sapiens chromosome 11 clone RP11-79904 map 11q14, WORKING
DEFINITION      DRAFT SEQUENCE, 32 unordered pieces.
ACCESSION      AP001320
VERSION      AP001320.2 GI:8117247
KEYWORDS      HTG; HTGS PHASE1; HTG; DRAFT.
SOURCE      Homo sapiens DNA, clone:RP11-79904.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 172830)
REFERENCE      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
AUTHORS      Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
              Homo sapiens 172,830 genomic DNA of 11q14
              Published Only in DataBase (2000)
              2 (bases 1 to 172830)
REFERENCE      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
AUTHORS      Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
              Direct Submission
              Submitted (01-MAR-2000) Masahira Hattori, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
              Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
              Japan (E-mail:hattori@gsc.riken.go.jp,
              URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
              Fax:81-42-778-9924)
              On May 30, 2000 this sequence version replaced gi:7209763.
COMMENT      ----- Genome Center
              Center: RIKEN Genomic Sciences Center (GSC)
              Center code: RIKEN
              Web site: http://hgp.gsc.riken.go.jp/
              Contact: hattori@gsc.riken.go.jp
              ----- Project Information
              Center project name: HumDraft11
              Center clone name: RP11-79904
              ----- Summary Statistics
              Sequencing vector: PCR products; 100% of reads
              Chemistry: Dye-terminator ET-amersham; 100% of reads
              Assembly program: Phrap; version 0.990329
              Consensus quality: 14586 bases at least Q40
              Consensus quality: 159404 bases at least Q30
              Consensus quality: 166388 bases at least Q20
              Insert size: 169730; sum-of-contigs
              Quality coverage: 4.17x in Q20 bases; sum-of-contigs
              -----
NOTE: This is a 'working draft' sequence. It currently consists of
32 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
              1
              15732 contig of 15732 bp in length
              15833 contig of 18269 bp in length
              34101 contig of 13278 bp in length
              47479 contig of 10177 bp in length
              57756 contig of 11570 bp in length
              69426 contig of 10859 bp in length
              80185 contig of 8704 bp in length
              88989 contig of 10273 bp in length
              93362 contig of 100 bp
              99463 contig of 3971 bp in length
              103434 contig of 100 bp
              103534 contig of 5531 bp in length
              109064 contig of 100 bp
              109164 contig of 100 bp
              10965 contig of 4691 bp in length
              109165 contig of 4691 bp in length
              113856 contig of 100 bp
              113956 contig of 5185 bp in length
              119140 contig of 100 bp
              119141 contig of 100 bp
              119241 contig of 5421 bp in length
              124662 contig of 100 bp
              124761 contig of 100 bp
              124762 contig of 2061 bp in length
              126822 contig of 100 bp
              126823 contig of 100 bp
              126923 contig of 4415 bp in length
              131337 contig of 100 bp
              131338 contig of 100 bp
              131438 contig of 3641 bp in length
              135078 contig of 100 bp
              135079 contig of 100 bp
              135178 contig of 100 bp
              135179 contig of 2956 bp in length
              138134 contig of 100 bp
              138135 contig of 100 bp
              138235 contig of 3889 bp in length
              142123 contig of 100 bp
              142124 contig of 100 bp
              142224 contig of 2496 bp in length
              144719 contig of 100 bp
              144720 contig of 100 bp
              144819 contig of 100 bp
              144820 contig of 4211 bp in length
              149031 contig of 100 bp
              149032 contig of 100 bp
              149131 contig of 2004 bp in length
              151134 contig of 100 bp
              151135 contig of 100 bp
              151235 contig of 2517 bp in length
              153751 contig of 100 bp
              153752 contig of 100 bp
              153851 contig of 2653 bp in length
              155604 contig of 100 bp
              156505 contig of 2979 bp in length
              159583 contig of 100 bp
              159684 contig of 100 bp
              159684 contig of 100 bp
              161669 contig of 1986 bp in length
              161670 contig of 100 bp
              161670 contig of 100 bp
              161770 contig of 2366 bp in length
              164136 contig of 100 bp
              164235 contig of 100 bp
              164236 contig of 1619 bp in length

```

* 165955 165954: gap of 100 bp
 * 165955 167495: contig of 1541 bp in length
 * 167496 167595: gap of 100 bp
 * 167596 168771: contig of 1176 bp in length
 * 168772 168871: gap of 100 bp
 * 168872 170322: contig of 1451 bp in length
 * 170323 170422: gap of 100 bp
 * 170423 171562: contig of 1140 bp in length
 * 171563 171662: gap of 100 bp
 * 171663 172830: contig of 1168 bp in length.

FEATURES

source
 1. .172830
 /location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosomes="11"
 /map="11q14"
 /clone="RP11-79904"

misc_feature
 1. .15732
 /note="assembly_fragment"
 misc_feature
 15933. .34101
 /note="assembly_fragment"
 misc_feature
 34202. .47479
 /note="assembly_fragment"
 misc_feature
 47580. .57756
 /note="assembly_fragment"
 misc_feature
 57857. .69426
 /note="assembly_fragment"
 misc_feature
 69527. .80185
 /note="assembly_fragment"
 misc_feature
 80286. .88989
 /note="assembly_fragment"
 misc_feature
 89090. .99362
 /note="assembly_fragment"
 misc_feature
 9463. .103433
 /note="assembly_fragment"
 misc_feature
 103534. .109064
 /note="assembly_fragment"
 misc_feature
 109165. .113855
 /note="assembly_fragment"
 misc_feature
 113956. .119140
 /note="assembly_fragment"
 misc_feature
 119241. .124661
 /note="assembly_fragment"
 misc_feature
 124762. .126822
 /note="assembly_fragment"
 misc_feature
 126923. .131337
 /note="assembly_fragment"
 misc_feature
 131438. .135078
 /note="assembly_fragment"
 misc_feature
 135179. .138134
 /note="assembly_fragment"
 misc_feature
 138235. .142123
 /note="assembly_fragment"
 misc_feature
 142224. .144719
 /note="assembly_fragment"

Query Match 78.0%; Score 175.6; DB 2; Length 172830;
 Best Local Similarity 88.9%; Pred. No. 6e-30;
 Matches 201; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 1 TGATGGTAGTCTTTTCAGGCATATAAATTTGAATAAATATGAGGCTCCATGATAGCT 60
 Db 28909 TGATGGTAGTCTTTTCAGGCATATAAATTTGAATAAATATGAGGCTCCATGATAGCT 28968
 QY 61 ATATTGGTTTACCTTCAGAGATATTTAGTTTCACTCAGGTTTTTCAAGCT-ACGCT 119
 Db 28969 ATATTGGTTTACCTTCAGAGATATTTAGTTTCACTCAGGTTTTTCAAGAGCTAACCT 29028
 QY 120 GTCCCCCAAAAAACGAAACAAAAACAAACACCTTTTAAAGAGTTGATGGCTACTCA 179
 Db 29029 GTCCCCCAAAATACGAGNACCAACCCCTTAAACACCCCTTTTAAAGAGTTGATGGCTAATCA 29088
 QY 180 TTTGATGTCGCTCCTCTGCTGGAATCAATTAGGAATTTTTTTTTTTT 225

Db 29089 TTGATTTCCCTCCTGCTGCTGCATCACATTAGGATTTTTTCTTTT 29134

RESULT 5

AC023384

LOCUS

DEFINITION

Homo sapiens chromosome 11 clone RP11-589112 map 11, LOW-PASS

SEQUENCE SAMPLING.

AC023384

AC023384

HTG; HTGS_PHASE0.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

REFERENCE

1 (bases 1 to 75002)

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC023384 75002 bp DNA linear HTG 13-JUL-2000
 Homo sapiens chromosome 11 clone RP11-589112 map 11, LOW-PASS
 SEQUENCE SAMPLING.

AC023384.2 GI:9144035

AC023384

HTG; HTGS_PHASE0.

Homo sapiens

Homo sapiens

REFERENCE

1 (bases 1 to 75002)

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

1674 2455: contig of 782 bp in length
2456 2555: gap of 100 bp
2556 3328: contig of 773 bp in length
3329 3428: gap of 100 bp
3429 4179: contig of 751 bp in length
4180 4279: gap of 100 bp
4280 5022: contig of 743 bp in length
5023 5122: gap of 100 bp
5123 5855: contig of 743 bp in length
5856 5965: gap of 100 bp
5966 6703: contig of 738 bp in length
6704 6803: gap of 100 bp
6804 7554: contig of 751 bp in length
7555 7654: gap of 100 bp
7655 8408: contig of 754 bp in length
8409 8508: gap of 100 bp
8509 9254: contig of 746 bp in length
9255 9354: gap of 100 bp
9355 10118: contig of 764 bp in length
10119 10218: gap of 100 bp
10219 10978: contig of 760 bp in length
10979 11078: gap of 100 bp
11079 11850: contig of 772 bp in length
11851 11950: gap of 100 bp
11951 12712: contig of 762 bp in length
12713 12812: gap of 100 bp
12813 13553: contig of 741 bp in length
13554 13653: gap of 100 bp
13654 14392: contig of 739 bp in length
14393 14492: gap of 100 bp
14493 15244: contig of 752 bp in length
15245 15344: gap of 100 bp
15345 16095: contig of 751 bp in length
16096 16195: gap of 100 bp
16196 16927: contig of 732 bp in length
16928 17027: gap of 100 bp
17028 17766: contig of 739 bp in length
17767 17866: gap of 100 bp
17867 18615: contig of 749 bp in length
18616 18715: gap of 100 bp
18716 19489: contig of 774 bp in length
19490 19589: gap of 100 bp
19590 20344: contig of 755 bp in length
20345 20444: gap of 100 bp
20445 21205: contig of 761 bp in length
21206 21305: gap of 100 bp
21306 22089: contig of 784 bp in length
22090 22189: gap of 100 bp
22190 22972: contig of 783 bp in length
22973 23072: gap of 100 bp
23073 23826: contig of 754 bp in length
23827 23926: gap of 100 bp
23927 24674: contig of 748 bp in length
24675 24774: gap of 100 bp
24775 25507: contig of 733 bp in length
25508 25607: gap of 100 bp
25608 26347: contig of 740 bp in length
26348 26447: gap of 100 bp
26448 27202: contig of 755 bp in length
27203 27302: gap of 100 bp
27303 28052: contig of 750 bp in length
28053 28152: gap of 100 bp
28153 28909: contig of 757 bp in length
28910 29009: gap of 100 bp
29010 29741: contig of 732 bp in length
29742 29841: gap of 100 bp
29842 30624: contig of 783 bp in length
30625 30724: gap of 100 bp
30725 31497: contig of 773 bp in length
31498 31597: gap of 100 bp
31598 32373: contig of 776 bp in length
32374 32473: gap of 100 bp
32474 33222: contig of 749 bp in length

33223 33322: gap of 100 bp
33323 34072: contig of 750 bp in length
34073 34172: gap of 100 bp
34173 34932: contig of 760 bp in length
34933 35032: gap of 100 bp
35033 35791: contig of 759 bp in length
35792 35891: gap of 100 bp
35892 36636: contig of 745 bp in length
36637 36736: gap of 100 bp
36737 37468: contig of 732 bp in length
37469 37568: gap of 100 bp
37569 38328: contig of 760 bp in length
38329 38428: gap of 100 bp
38429 39191: contig of 763 bp in length
39192 39291: gap of 100 bp
39292 40036: contig of 745 bp in length
40037 40136: gap of 100 bp
40137 40895: contig of 759 bp in length
40896 40995: gap of 100 bp
40996 41729: contig of 734 bp in length
41730 41829: gap of 100 bp
41830 42577: contig of 748 bp in length
42578 42677: gap of 100 bp
42678 43414: contig of 737 bp in length
43415 43514: gap of 100 bp
43515 44267: contig of 753 bp in length
44268 44367: gap of 100 bp
44368 45123: contig of 756 bp in length
45124 45223: gap of 100 bp
45224 45973: contig of 750 bp in length
45974 46073: gap of 100 bp
46074 46816: contig of 743 bp in length
46817 46916: gap of 100 bp
46917 47681: contig of 765 bp in length
47682 47781: gap of 100 bp
47782 48544: contig of 763 bp in length
48545 48644: gap of 100 bp
48645 49410: contig of 766 bp in length
49411 49510: gap of 100 bp
49511 50247: contig of 737 bp in length
50248 50347: gap of 100 bp
50348 51067: contig of 720 bp in length
51068 51167: gap of 100 bp
51168 51915: contig of 748 bp in length
51916 52015: gap of 100 bp
52016 52745: contig of 730 bp in length
52746 52845: gap of 100 bp
52846 53584: contig of 739 bp in length
53585 53684: gap of 100 bp
53685 54427: contig of 743 bp in length
54428 54527: gap of 100 bp
54528 55292: contig of 765 bp in length
55293 55392: gap of 100 bp
55393 56164: contig of 772 bp in length
56165 56264: gap of 100 bp
56265 56985: contig of 721 bp in length
56986 57085: gap of 100 bp
57086 57829: contig of 744 bp in length
57830 57929: gap of 100 bp
57930 58677: contig of 748 bp in length

Query Match 57.1%; Score 128.4; DB 2; Length 75002;
Best Local Similarity 98.6%; Pred. No. 2,3e-19;

Matches 140; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 84 ATATTAGTTTCACCTCAGGTTTTTCAAGCTACGCTGCCCAAAACGAAACAAAC 143

DB 35932 ATATTAGTTTCACCTCAGGTTTTTCAAGCTACGCTGCCCAAAACGAAAC 35990

QY 144 AAAAAACAACCTTTTAAAGAGTTTATGCTACTCATTTGATCTGCTCTCTGCTGAAT 203

DB 35991 AAAAAACAACCTTTTAAAGAGTTTATGCTACTCATTTGATCTGCTCTCTGCTGAAT 36050

AC094553 192454 bp DNA linear HTG 10-JUL-2000
Rattus norvegicus clone CH230-4p5, *** SEQUENCING IN PROGRESS ***
56 unordered pieces.
AC094553
AC094553.6 GI:21716350
HTG; HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Mammalia; Metazoa; Chordata; Vertebrata; Euteleostomi;
Eumaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 192454)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrechts,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,K., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chaves,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.J., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamillon,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoguea,M., Hollway,C., Hollins,B.,
Homi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureahi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Petrels,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vaquez,L., Vera,V., Vallalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 192454)
Worley,K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 192454)
Worley,K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 9, 2002 this sequence version replaced gi:19874113.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GAWO
Center clone name: CH230-4P5
----- Summary Statistics

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 138018 bases at least Q40
Consensus quality: 144061 bases at least Q30
Consensus quality: 149644 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1247: contig of 1247 bp in length
* 1248: gap of unknown length
* 1348: contig of 1104 bp in length
* 2451: gap of unknown length
* 2551: contig of 1203 bp in length
* 2552: gap of unknown length
* 3755: contig of 1056 bp in length
* 3855: gap of unknown length
* 4911: contig of 1058 bp in length
* 5011: gap of unknown length
* 6069: contig of 1058 bp in length
* 6169: gap of unknown length
* 7868: contig of 1699 bp in length
* 7968: gap of unknown length
* 9574: contig of 1606 bp in length
* 9674: gap of unknown length
* 11193: contig of 1520 bp in length
* 11194: gap of unknown length
* 12334: contig of 1041 bp in length
* 12335: gap of unknown length
* 13435: contig of 1008 bp in length
* 13443: gap of unknown length
* 13543: contig of 1036 bp in length
* 14578: gap of unknown length
* 14579: contig of 1216 bp in length
* 15894: gap of unknown length
* 15895: contig of 1048 bp in length
* 15985: gap of unknown length
* 17043: contig of 1224 bp in length
* 17143: gap of unknown length
* 18367: contig of 1774 bp in length
* 18467: gap of unknown length
* 20241: contig of 1340 bp in length
* 20341: gap of unknown length
* 21681: contig of 1242 bp in length
* 21781: gap of unknown length
* 23023: contig of 1335 bp in length
* 23123: gap of unknown length
* 24457: contig of 1369 bp in length
* 24558: gap of unknown length
* 25927: contig of 1455 bp in length
* 26037: gap of unknown length
* 27481: contig of 1019 bp in length
* 27581: gap of unknown length
* 28600: contig of 1019 bp in length
* 28700: gap of unknown length
* 28701: contig of 2505 bp in length
* 31205: gap of unknown length
* 31305: contig of 2480 bp in length
* 33785: gap of unknown length
* 33885: contig of 1709 bp in length
* 33886: gap of unknown length
* 35694: contig of 2496 bp in length
* 35695: gap of unknown length
* 38190: contig of 2394 bp in length
* 38191: gap of unknown length
* 40684: contig of 2249 bp in length
* 40784: gap of unknown length
* 43033: contig of 2249 bp in length

43034 43133: gap of unknown length
* 43134: contig of 2006 bp in length
* 45140: gap of unknown length
* 45240: contig of 2074 bp in length
* 47314: gap of unknown length
* 47414: contig of 2737 bp in length
* 50151: gap of unknown length
* 50251: contig of 2103 bp in length
* 52354: gap of unknown length
* 52454: contig of 3358 bp in length
* 55811: gap of unknown length
* 55911: contig of 3286 bp in length
* 59197: gap of unknown length
* 59297: contig of 3040 bp in length
* 62337: gap of unknown length
* 62338: contig of 3943 bp in length
* 64338: gap of unknown length
* 66381: contig of 3740 bp in length
* 66481: gap of unknown length
* 70221: contig of 3546 bp in length
* 70321: gap of unknown length
* 73867: contig of 2749 bp in length
* 73967: gap of unknown length
* 76716: contig of 4117 bp in length
* 76816: gap of unknown length
* 80933: contig of 3364 bp in length
* 81033: gap of unknown length
* 84397: contig of 4147 bp in length
* 84497: gap of unknown length
* 88644: contig of 3994 bp in length
* 88744: gap of unknown length
* 92738: contig of 2491 bp in length
* 92838: gap of unknown length
* 95329: contig of 5671 bp in length
* 95429: gap of unknown length
* 101100: contig of 4831 bp in length
* 101200: gap of unknown length
* 106031: contig of 5026 bp in length
* 106131: gap of unknown length
* 111157: contig of 4641 bp in length
* 111257: gap of unknown length
* 115898: contig of 4959 bp in length
* 115998: gap of unknown length
* 120957: contig of 6864 bp in length
* 121057: gap of unknown length
* 127921: contig of 5996 bp in length
* 128021: gap of unknown length
* 134017: contig of 9289 bp in length
* 134117: gap of unknown length
* 143406: contig of 7766 bp in length
* 143506: gap of unknown length
* 151272: contig of 11271 bp in length
* 151272: gap of unknown length
* 162643: contig of 10440 bp in length
* 162743: gap of unknown length
* 173182: contig of 10440 bp in length

Query Match 20.5%; Score 46.2; DB 2; Length 192454;
Best Local Similarity 54.4%; Pred. No. 0.42; Indels 0; Gaps 0;
Matches 93; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 24 AAAATTGAAATAAATATGAGGCTCCATGATATGCTATATGTTTACCTTCAGAAGA 83
Db 113461 AAAATTGCTCAATTTCTTTGTATACAAAGATTGAGGCTTTATTTTCCCATAGAGAA 113520
QY 84 ATATTTAGTTTCACTCAGGTTTTTCAAGCTAGCTGCTCCCCAAAAAGAACAAAC 143
Db 113521 TCGTTGAATTTCACTAAGAAATTTTCAAGTACATTTAGTCTGCTAAAAAAGAAAG 113580
QY 144 AAAAAAACACCTTTTAAAGAGTTGATGCTACTCATTTGATCTGCTCTCT 194
Db 113581 AAAAGAGAGAGAAATTTCCAGTTGATTTACTAATAAGATTTCTCTT 113631

RESULT 8

U67577/C	9971 bp	DNA	linear	BCT 28-JAN-1998
LOCUS				
DEFINITION	Methanococcus jannaschii section 119 of 150 of the complete genome.			
ACCESSION	U67577	L77117		
VERSION	U67577.1	GI:2826400		
KEYWORDS				
SOURCE	Methanocaldococcus jannaschii.			
ORGANISM	Methanocaldococcus jannaschii			
REFERENCE	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanocaldococcaeae; Methanocaldococcus.			
AUTHORS	1 (bases 1 to 9971) Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A., Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G., Merrick,J.M., Glodek,A., Scott,J.D., Geoghagen,N.S., Weidman,J.F., Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Kelley,J.M., Peterson,J.D., Sadow,P.W., Hanna,M.C., Cotton,M.D., Hurst,M.A., Roberts,K.M., Kaine,B., Borodovsky,M., Klenk,H.P., Fraser,C.M., Smith,H.O., Woese,C.R. and Venter,J.C.			
ITILE	Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii			
JOURNAL	Science 273 (5278), 1058-1073 (1996)			
MEDLINE	96337999			
PUBMED	8688087			
REFERENCE	2 (bases 1 to 9971) Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A., Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J.-F., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G., Merrick,J.M., Glodek,A., Scott,J.D., Geoghagen,N.S., Weidman,J.F., Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Kelley,J.M., Peterson,J.D., Sadow,P.W., Hanna,M.C., Cotton,M.D., Hurst,M.A., Roberts,K.M., Kaine,B., Borodovsky,M., Klenk,H.P., Fraser,C.M., Smith,H.O., Woese,C.R. and Venter,J.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-AUG-1996) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
COMMENT	On Jan 30, 1998 this sequence version replaced gi:1592013.			
FEATURES	Location/Qualifiers			
source	1..9971 /organism="Methanocaldococcus jannaschii" /db_xref="taxon:2190" 87..554 /gene="MJ1369" 87..554 /gene="MJ1369" /note="hypothetical protein; identified by GeneMark; putative" /codon_start=1 /transl_table=11 /product="M. jannaschii predicted coding region MJ1369" /protein_id="AAB99386.1" /db_xref="GI:1592014" /translation="MDIIVTWLFPSSNIRNIEICYNHMGFWDRNAGEKQKNWR SFIRKNQIKPFDVAVQIAKGEIHAIGVETIYDDQPIWDNEINLKNVTFPMRV SFSVIIISKEAVIKRPIKIQDYIGLGELEHDFNEILKAPQKFGMSIK"			
gene	584..1570 /gene="MJ1370" 584..1570 /gene="MJ1370" /note="hypothetical protein; identified by GeneMark; putative"			
CDS	/transl_table=11 /product="M. jannaschii predicted coding region MJ1370" /protein_id="AAB99387.1" /db_xref="GI:1592015" /translation="MTVGDHGTSGITTCIKDNKKIIFKLKRTLKEKSYLEELEK HISLEDIALTYSMGDGINKILPIEKVKNRGVLISIEGAKVGGTKVYDEIKESG LPAAVYIPLHGRFECLEDERFRALYSHASPEKVSIAIYAYKLFQFNDFVLSISNTV TLLIKGKIFGGDACIAGIMLHGFIDEMINDIDAGKITANEAFSKAGAVKIAKLY KGVENTKEEIIINNFYDENCRLAVDSLLILSVSMEINSLPLLDKKNRRVVLASIGTL			
gene	3680..4480 /gene="MJ1374" 3680..4480 /gene="MJ1374" /note="similar to GB:AE000666 percent identity: 42.73; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAB99381.1" /db_xref="GI:1592019" /translation="MRVEIIFLFCGGGRWATITQKATGGFRIHTNLRMHVDPGPGA IURLNELKISPMRTNALFISCHPDHYTDEIIVBAITQCMTKKRGVFLGSLSVSGF GEYEVVSKYHOSKLEEVRLYPGDSAEIYDTTIKATHTKGDGPGFGLSLTYGDI GYTSTERTIPQIIEFDGVRILLIANIVRKKNRIKGLHCLSDAIDLINSNKKPELLI MNHMGKMTNPQIEAEYSQNGIEIVPARLGLKVELLNGRYQILIK"			
CDS	complement(4473..5720) /gene="MJ1375" complement(4473..5720) /gene="MJ1375" /note="similar to GB:D26185 SP:P37555 PID:467446 GB:AL009126 percent identity: 23.61; identified by sequence similarity; putative"			
gene	3124..3612 /gene="MJ1373" 3124..3612 /gene="MJ1373" /note="hypothetical protein; identified by GeneMark; putative" /codon_start=1 /transl_table=11 /product="M. jannaschii predicted coding region MJ1373" /protein_id="AAB99388.1" /db_xref="GI:1592018" /translation="MGILGKIKNKLKSGIKRVILDFYRLDNSGVULKIYEKILEBAID KNLSPHVAIMDGNRAAEIYKORYYGHYGAERVRLRWARDLGNVVTLYAFS TENFRPKEVDLMELPEKFEIADDEIHYEVVRVRAIGRLNLPKNVQAKIAYA LERTKNYKFFVIAIAYGQGEIIDAIVKKAIEKVRGEIPEIDIKELIDKHLTYAN LFPNPDLIIRTSGBERISNFIWQSSYSELVFCDIYWFLFRVDFLRAVRDQRRQR REGK"			
CDS	2285..3127 /gene="MJ1372" 2285..3127 /gene="MJ1372" /note="similar to GB:L42023 SP:P44938 PID:1006041 PID:1221025 PID:1205167 percent identity: 47.86; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAB99380.1" /db_xref="GI:1592017" /translation="MGILGKIKNKLKSGIKRVILDFYRLDNSGVULKIYEKILEBAID KNLSPHVAIMDGNRAAEIYKORYYGHYGAERVRLRWARDLGNVVTLYAFS TENFRPKEVDLMELPEKFEIADDEIHYEVVRVRAIGRLNLPKNVQAKIAYA LERTKNYKFFVIAIAYGQGEIIDAIVKKAIEKVRGEIPEIDIKELIDKHLTYAN LFPNPDLIIRTSGBERISNFIWQSSYSELVFCDIYWFLFRVDFLRAVRDQRRQR REGK"			
gene	1567..2301 /gene="MJ1371" 1567..2301 /gene="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="WIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLLVDKLL DGEIDGAVRGSSSKVLIDYLRERICKFYRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNFNLKAVILSGRGLDGRNKKVVDETIYAEAEIIEVHFK GNVDIITHNGILIEEVLKCGYNIIAVDGTIGNLIFRLGLICKIPGYGAVILSDKNVN FIDTSRNNWERYNAIKFLIGGDFG"			
CDS	2285..3127 /gene="MJ1372" 2285..3127 /gene="MJ1372" /note="similar to GB:L42023 SP:P44938 PID:1006041 PID:1221025 PID:1205167 percent identity: 47.86; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAB99380.1" /db_xref="GI:1592017" /translation="MGILGKIKNKLKSGIKRVILDFYRLDNSGVULKIYEKILEBAID KNLSPHVAIMDGNRAAEIYKORYYGHYGAERVRLRWARDLGNVVTLYAFS TENFRPKEVDLMELPEKFEIADDEIHYEVVRVRAIGRLNLPKNVQAKIAYA LERTKNYKFFVIAIAYGQGEIIDAIVKKAIEKVRGEIPEIDIKELIDKHLTYAN LFPNPDLIIRTSGBERISNFIWQSSYSELVFCDIYWFLFRVDFLRAVRDQRRQR REGK"			
gene	3680..4480 /gene="MJ1374" 3680..4480 /gene="MJ1374" /note="similar to GB:AE000666 percent identity: 42.73; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAB99381.1" /db_xref="GI:1592019" /translation="MRVEIIFLFCGGGRWATITQKATGGFRIHTNLRMHVDPGPGA IURLNELKISPMRTNALFISCHPDHYTDEIIVBAITQCMTKKRGVFLGSLSVSGF GEYEVVSKYHOSKLEEVRLYPGDSAEIYDTTIKATHTKGDGPGFGLSLTYGDI GYTSTERTIPQIIEFDGVRILLIANIVRKKNRIKGLHCLSDAIDLINSNKKPELLI MNHMGKMTNPQIEAEYSQNGIEIVPARLGLKVELLNGRYQILIK"			
CDS	complement(4473..5720) /gene="MJ1375" complement(4473..5720) /gene="MJ1375" /note="similar to GB:D26185 SP:P37555 PID:467446 GB:AL009126 percent identity: 23.61; identified by sequence similarity; putative"			

```
/codon_start=1
/transl_table=11
/product="spovB isolog"
/protein_id="AA999382.1"
/db_xref="GI:2826401"
/translation="MGLKDSIYILMSNLVSKGMAYLFVYFITAFLGLTBAFGILKGLM
PIADTUTPFSSGIPAPIAFLAEKEVDINKYIPILYMLLSVVGFFLTPYIKVIL
GHLVLPNTLYFAVGLCVVASTVIAFSRGIQLGLMKYLSLWTVEYTKVLVFI
LTLVGLFGLSLISLAYLVGGIFGLYLYALKKLGKDFPKLIDIKNTVKNFSPNL
DILRYPIPALTSSTRLFGDINDINIVMSIMGFWGSIYGYSSLSRGLPMFASVSI
PLPRLSKTKDLSLKEGIIQNTIFSSIFVIGCLFPEIPLEIAFFTAANPEGLCLURI
LAISSFMSYITLISALOGLYAKISFYILFGLVLNLIILNLVNAVYGVGGSLAT
LITSVIFLIGFAILRIKKHII"
gene 5807.6544
CDS /gene="MJ1376"
5807.6544
/notes="MJ1376"
PID:1221464 PID:1205570 percent identity: 40.10;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cell division protein FtsZ"
/protein_id="AA899383.1"
/db_xref="GI:1592021"
/translations="MGRKDRVLRQRDPFYKLAKKLYRSPASFKLMQNEKFNVI
KPKVILDLGCAGGMOVAREIVGKGFVIGIDLPVFPPEYDYNVAIKGDTLEEN
LNKIRELPNDEKRVVIVSDASPNISGWIDHAKSIDLVTTALQIATEMLKERNF
VAKVFGVDNEDLVNIVKRYFEKYITKPOASRKEAEVVIKRYTKGKKWBEEDKIK
RKKVFNEDNELLAKIKIRKLKSKK"
gene 5807.6544
CDS /gene="MJ1377"
5807.6544
/notes="MJ1377"
PID:1051202 SP:Q50543 percent identity:
51.36; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glutaryl-tRNA synthetase (gltX)"
/protein_id="AA899384.1"
/db_xref="GI:1592022"
/translations="MEKILPIALRNAIKYNGKANPKAVIGILFSENPEYRSKAEWM
PIVKEVVEVKNLSDBEIKKLELGEVDKKEKGLLEPNVKDKVMVRFAPNPSG
PLIGHARAALVNDYVKKYGGKILRLLEDTPKVLPEAYDVKEDLWGLVQKDEEV
VIOSDRILEYVGRKLIMGHAYVDCNPEERELRNKGVPCKCDRAEDLNELEWE
KMLGENVAVRKIDIKHNSIRDFPIFRVEKTPHPTGDKYCYVPLMPSVPVD
DHLGMTHLVGRKDHVNTKQYIYKFGWEMPEFIHYGKILKIEDIVLSTSMYKGI
KEGLYSQDDVRLGTLRALRRRIGKPEAIYEIMKRIQKQADVFSWENLYAINKELI
DKDARFFVFNPKKLIIEGAERKVLKRMHPDRPFGERELIFDGEVYVVGDELEN
KMYRLMELFNVVEKVDIDIALKYHSDDFKIRKKNKAKIHWIPVKDSVKVKVLMPPDG
EIKGEFAEKDFAKVEVDIIQFERFGEVRIDKNDGDFVCCYVHR"
gene 8402.9850
CDS /gene="MJ1378"
8402.9850
/genes="MJ1378"
/genes="MJ1378"
Query Match 20.18; Score 45.2; DB 1; Length 9971;
Best Local Similarity 51.04; Pred. No. 1.2;
Matches 107; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 7 TAAGTTCTTTTCAGGCATATAAATTTGAAATATATGAGGCTCATGATGCTATATTG 66
DB 2481 TAAATCTCTGAGCCCTTCTATTCCATCATATTATATAGCTACATGTTTGGTAAATTG 2422
QY 67 GTTTTACCTTCAGAGAAATATTAGTTTCTACTCAGGTTTTCAAAGCTACGCTGCCCC 126
DB 2421 TCTTTATCAATAGCTCTTCTAAATCTTCTCGTAAATTTTAAACCTCGGAGTGTCT 2362
QY 127 AAAAAAGCAACAAACAAACAAACACCTTTTAAAGTTGATGGTACTCATTTGATC 186
DB 2361 AAAAACTATAAAAAAATCAATATTACTCTTTTCCAACTACTCTTAAATTTGTTTATC 2302
QY 187 TGCCTCTCTGCTGAATCAATAGGAATTT 216
```



```

source      1. .200110
            /organism="Mus musculus"
            /db xref="taxon:10090"
            /clone="RP23-431G18"
BASE COUNT  60329 a 37978 c 38823 g 62980 t
ORIGIN
Query Match      19.5%; Score 43.8; DB 10; Length 200110;
Best Local Similarity 55.6%; Pred. No. 1.4;
Matches 84; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 75 TTCAAGAAATATTAGTTTCATCAGGTTTTCACAAAGCTACGCTGCTCCCAAAAACG 134
    |||||
Db 105570 TACAGAGTGAAGTCCAGGAGCAGCGGTACACAGAGAACCTGTCTCAAAAAA 105511
    |||||
QY 135 ARAACAACAACAAACAACTTTTAAAGATGTAGGCTACTCAATTTGATCTGCTCCT 194
    |||||
Db 105510 AAAAAAACAACAAACAACTTTTAAAGATGTAGGCTACTCAATTTGATCTGCTCCT 194
    |||||
195 CTGCTGAATCAATTAGGAATTTTTTTTTTTT 225
105450 TTATTTTACATTAGGTGTTATGTGTTT 105420

RESULT 11
AP005406/c      128526 bp DNA linear HTG 13-JUN-2002
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 8 clone B1147B12,
*** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION
AP005406
VERSION
AP005406.1 GI:21396516
KEYWORDS
HTG; HTGS PHASE2.
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:B1147B12.
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 Sasaki,T., Matsumoto,T. and Katayose,Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
clone:B1147B12
Published Only in Database (2002)
JOURNAL
2 (bases 1 to 128526)
REFERENCE
Sasaki,T., Matsumoto,T. and Katayose,Y.
Direct Submission
Submitted (12-JUN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .128526
   Location/Qualifiers
   /organism="Oryza sativa (japonica cultivar-group)"
   /cultivar="Nipponbare"
   /db xref="taxon:39947"
   /chromosome="8"
   /clone="B1147B12"
BASE COUNT  37912 a 25748 c 26754 g 37999 t 113 others
ORIGIN
Query Match      18.9%; Score 42.6; DB 2; Length 128526;
Best Local Similarity 51.9%; Pred. No. 2.9;

```

```

Matches 96; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 30 TCAATAAATATGAGGCTCCATGATATGCTATATTTGTTTACCTTCAGAGAATATTT 89
    |||||
Db 35653 TGAATAGCTGTGATATCCAGGAAATCGTCATCATTTTCAAGACATTTGGATGATAG 35594
    |||||
QY 90 AGTTTCACTCAGGTTTTTCAAGCTACGCTGCTCCCAAAAACGAAACCAAAAAA 149
    |||||
Db 35593 AATCCAGAGAGCTCAAGTTTTCATTTGCTGACCCCAAAAAGAAACCAACAAA 35534
    |||||
QY 150 ACAACCTTTTAAAGAGTTGATGCTACTCATTTGATCTGCTCTGCTGAATTA 209
    |||||
Db 35533 GCATACATGAAAAATGTCAATGATAAGTTAATAAGATGGTACAAAGCATATCCAAACCA 35474
    |||||
QY 210 GGAAT 214
    |||||
Db 35473 CAATT 35469
    |||||

RESULT 12
AC006281/c      199551 bp DNA linear HTG 12-AUG-2000
LOCUS
DEFINITION
Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 2 unordered pieces.
ACCESSION
AC006281
VERSION
AC006281.8 GI:97977738
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
Plasmodium falciparum.
ORGANISM
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 199551)
Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
Kurdi,O.B., Conway,A.B. and Davis,R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
2 (bases 1 to 199551)
Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B. and Davis,R.W.
Direct Submission
Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo-Alto, CA
94304, USA
COMMENT
On Aug 12, 2000 this sequence version replaced gi:8810457.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 41526: contig of 41526 bp in length
* 41527 41726: gap of unknown length
* 41727 199551: contig of 157825 bp in length.
FEATURES
source
1. .199551
   Location/Qualifiers
   /organism="Plasmodium falciparum"
   /db xref="taxon:5833"
   /chromosome="12"
   /clone="PFYACB8-95"
   /clone="3D7"
BASE COUNT  83197 a 17543 c 19759 g 78852 t 200 others
ORIGIN
Query Match      18.8%; Score 42.4; DB 2; Length 199551;
Best Local Similarity 50.5%; Pred. No. 2.9;
Matches 103; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 22 ATAAATTTGAAATAAATATGAGGCTCCATGATGCTATATTTGTTTACCTTCAGAA 81
    |||||
Db 22946 AAAAAAATAAATAAACAACATATTTTATATATATATATATATATGTAAG 22887
    |||||
QY 82 GAATATTTAGTTTCTACTCAGGTTTTTCAAGCTACGCTCTCCCAAAAACGAAACAA 141
    |||||

```

```

Db 22886 TATTATTATAAATCAATTTTCATATAATATATATAAAAAAAAAAAAAAAAAAAAA 22827
OY 142 ACAAATAAAACCTTTTAAAGTGTGATGGCTACTCATTTGATCGCTCTCTCTGTA 201
Db 22826 AAAAAAAAAAAAAAATGTTATAACACCATCGTTCTCTCTCTCTCTCTCTCTCTTT 22767
OY 202 ATCAATAGCAATTTTTTTTTTTT 225
Db 22766 TTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 22743

RESULT 13
AC011846
LOCUS Homo sapiens chromosome 15, clone RP11-346A8, complete sequence.
AC011846
AC011846.16 GI:21389294
HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 147760)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, clone RP11-346A8
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 147760)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckler,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Castie,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donean,B., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 147760)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicoll,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (10-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 147760)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicoll,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (10-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

TITLE
JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L2618

Center clone name: 346_A_8

Only the last 147.8 kb of this clone are being submitted.

The remainder overlaps accession number AC105129 [WICGR project L22971].

FEATURES

source

Location/Qualifiers

1. .147760

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosomes="15"

/maps="15"

/clone="RP11-346A8"

/clone_lib="RP11-11 Human Male BAC"

572..672

/rpt_family="MIR"

697..1138

/rpt_family="LIMB6"

1143..1340

/rpt_family="AluJo"

1357..1665

/rpt_family="AluSg"

1666..1699

/rpt_family="L2"

1771..2026

/rpt_family="AluJo"

2049..2069

/rpt_family="(TC)n"

2072..2096

/rpt_family="(CA)n"

2154..2185

/rpt_family="(TG)n"

complement(2458..2533)

/rpt_family="MIR3"

2943..3081

/rpt_family="MIR3"

complement(3473..3526)

/rpt_family="L2"

3584..3758

repeat_region


```

repeat_region /rpt_family="MIR"
complement(3767..4021)
repeat_region /rpt_family="AluSx"
complement(4075..4115)
repeat_region /rpt_family="AluSx"
complement(4737..4774)
repeat_region /rpt_family="MIR"
5062..5168
repeat_region /rpt_family="MIR"
complement(5262..5563)
repeat_region /rpt_family="AluY"
complement(6726..7022)
repeat_region /rpt_family="AluSx"
7049..7166
repeat_region /rpt_family="L2"
7215..7332
repeat_region /rpt_family="MIR"
complement(7387..7530)
repeat_region /rpt_family="L3"
8050..8168
repeat_region /rpt_family="MIR3"
complement(8169..8480)
repeat_region /rpt_family="AluDo"
8481..8537
repeat_region /rpt_family="MIR3"
complement(9062..9373)
repeat_region /rpt_family="AluSx"
9383..9601
repeat_region /rpt_family="L2"
9755..10149
repeat_region /rpt_family="L2"
11647..11688
repeat_region /rpt_family="(CA)n"
11960..11991
repeat_region /rpt_family="GC-rich"
12008..12028
repeat_region /rpt_family="(CCG)n"
12084..12161
repeat_region /rpt_family="G-rich"
complement(12722..12877)
repeat_region /rpt_family="MIR3"
13023..13086
repeat_region /rpt_family="L3"
13212..13282
repeat_region /rpt_family="L2"
13283..13537
repeat_region /rpt_family="L2"
13538..13667
repeat_region /rpt_family="FLAM_C"
13668..14227
repeat_region /rpt_family="L2"
complement(14228..14513)
repeat_region /rpt_family="AluSx"
14514..14525
repeat_region /rpt_family="L2"
complement(14526..14835)
repeat_region /rpt_family="AluSx"
14836..14999
repeat_region /rpt_family="L2"
complement(15356..15637)
repeat_region /rpt_family="L2"
15859..15991
repeat_region /rpt_family="MIR3"
16064..16363
repeat_region /rpt_family="AluSx"
16597..16909
repeat_region /rpt_family="AluJb"

```

Query Match 18.7%; Score 42; DB 9; Length 147760;
 Best Local Similarity 52.9%; Pred.No.3.8;
 Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

17 CAGGCATAAAATTTGAAATAAATATGAGGCTCCATGATATGCTATATGTTGGTTTACCTT 76

```

Db 107925 CACTAATAAAAAATCCATTTGGTACTATGCTCCACTCCACAGTATATAATTTATTCACCTC 107984
QY 77 CAGAGAATATTTAGTTTTCACTCAGGTTTTTCAAAGCTACGCTGTCCTCCCAAAAAACGAA 136
Db 107985 TGAGGTTAAATGAGGCTTCCCTCAGGCTTTTAAAGTTGAGCTCCGTTAAAAAAGAAAAA 108044
QY 137 ACAAAACAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATC 186
Db 108045 AAAAAAAGGCAAGTCATTTTCTTGGTCTCTCTTCTATTAATTTCTC 108094

RESULT 14
AC067745/c 156140 bp DNA linear PRI 10-MAY-2002
LOCUS Homo sapiens chromosome 10 clone RP11-417C21, complete sequence.
DEFINITION AC067745
ACCESSION AC067745
VERSION AC067745.7 GI:20522196
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 156140)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 156140)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 3 (bases 1 to 156140)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On May 10, 2002 this sequence version replaced gi:13811871.
FEATURES
source
1..156140
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-417C21"
/clone_lib="RPCI-11"
BASE COUNT 44203 a 38631 c 34570 g 38736 t
ORIGIN
Query Match 18.6%; Score 41.8; DB 9; Length 156140;
Best Local Similarity 50.2%; Pred.No.4.2;
Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 21 CATAAAATTTGAAATAAATATGAGGCTCCATGATATGCTATATGTTGGTTTACCTTCAGA 80
Db 139039 CGTTTAATTCGTTTATATCTAATACTACCTGATACATAATGCTAGTTTCTCTGT 138980
QY 81 AGAATATTTAGTTTCACTCAGGTTTTTCAAAGCTACGCTCCCAAAAAACGAAACAA 140
Db 138979 TTTATTTTGGGTTCTTCTGCTCTCTTTCCCTCTCTTATTAATTAAGAAATTTGT 138920
QY 141 AACAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATCTGCTCTCTGCTG 200
Db 138919 TCCATATTTCTTCTTTTAAATTTTGGTTTCATAGCTAGATCATATTCGTTC 138860
QY 201 AATCAATTAGGAATTTTTTTTTTTT 225
Db 138859 TCCTAGTGATTACTCTTTTTTTTTT 138835

RESULT 15
AC013420 215467 bp DNA linear HTG 27-AUG-2000
LOCUS AC013420

```



```

DEFINITION Homo sapiens clone RP11-619F23, WORKING DRAFT SEQUENCE, 19
AC013420
AC013420.13 GI:9929508
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 215467)
Muzny,D.M., Adams,C., Baile,M., Barbara,J., Blankenburg,K.,
Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Taney,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wabnah,M.,
Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 215467)
Worley,K.C.
Direct Submission
Submitted (10-NOV-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Aug 26, 2000 this sequence version replaced gi:9719539.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMMP
Center clone name: RP11-619F23
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 175806 bases at least Q40
Consensus quality: 195338 bases at least Q30
Consensus quality: 202378 bases at least Q20
Estimated insert size: 207191; sum-of-contigs estimation
Estimated insert size: 249405; agarose-gel estimation
Quality coverage: 2.7x in Q20 bases; agarose-gel estimation
Quality coverage: 3.3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 19 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 30720: contig of 30720 bp in length
* 30721 30820: gap of unknown length
* 30821 57766: contig of 26946 bp in length
* 57767 57866: gap of unknown length
* 57867 79135: contig of 21269 bp in length

```

```

* 79136 79235: gap of unknown length
* 79236 93871: contig of 14636 bp in length
* 93872 109687: contig of 15716 bp in length
* 109688 109787: gap of unknown length
* 109788 124632: contig of 14845 bp in length
* 124633 124732: gap of unknown length
* 124733 139466: contig of 14614 bp in length
* 139467 139447: gap of unknown length
* 139448 150488: contig of 11042 bp in length
* 150489 150588: gap of unknown length
* 150589 162913: contig of 12325 bp in length
* 162914 163013: gap of unknown length
* 163014 171860: contig of 8647 bp in length
* 171861 171761: gap of unknown length
* 171762 178556: contig of 6996 bp in length
* 178557 178556: gap of unknown length
* 178557 186909: contig of 8053 bp in length
* 186910 187009: gap of unknown length
* 187010 194538: contig of 7629 bp in length
* 194539 194738: gap of unknown length
* 194739 199508: contig of 4770 bp in length
* 199509 199608: gap of unknown length
* 199609 204158: contig of 4550 bp in length
* 204159 204258: gap of unknown length
* 204259 207133: contig of 2875 bp in length
* 207134 207233: gap of unknown length
* 207234 210635: contig of 3302 bp in length
* 210636 214187: gap of unknown length
* 214188 214287: contig of 3552 bp in length
* 214288 215467: contig of 1180 bp in length.
FEATURES             Location/Qualifiers
     source            1..215467
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="RP11-619F23"
BASE COUNT   58444 a 50354 c 49030 g 55779 t 1860 others
ORIGIN
Query Match      18.6%; Score 41.8; DB 2; Length 215467;
Best Local Similarity 50.2%; Pred. No. 3.9;
Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 21 CATAAATTTGAAATAAATATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGA 80
DB 190461 CGTTTAATTCGTTTATACCTTAATATACACCTGATACATATGTCAGTTTCTGCTCT 190520
QY 81 AGAATATTAGTTTTCACCTCAGGTTTTTCAAGCTACGCTGTCCTCCCAAAAACGAAACAA 140
DB 190521 TTTATTTTGGGTTTCTTTGCTCCTCTTTCCCTTTCTTATATTAAAGGAATTGT 190580
QY 141 AACAAAAAACAACTTTTAAAGCTTGTGCTACTCATTTGATCGCTCTCTCTGCTG 200
DB 190581 TCCATATTCTTCTTTTAAATTTTGTGTTTGGTTTCATAGCTGTAGATCATATTCTGTT 190640
QY 201 AATCAATTAGGAATTTTTTTTTTTT 225
DB 190641 TCCTAGTGATTAATCTTTTTTTTTT 190665

```

Search completed: June 7, 2003, 20:57:10
Job time : 1311 secs

THIS PAGE BLANK (USPTO)

Schnitzer, R.
091765223/page 1
seq ID 58 w/
inters

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:50 ; Search time 1687 Seconds
(without alignments)
3353.325 Million cell updates/sec

Title: US-09-765-231A-58

Perfect score: 225

Sequence: 1 tgaatggaagttgtttccagg.....attaggaatttttttttttt 225

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

al number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:*

1: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023 COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	225	100.0	225	30	US-09-765-231A-58			Sequence 58, Appl
2	225	100.0	225	61	US-60-176-523-58			Sequence 58, Appl
3	225	100.0	320	18	US-09-495-050A-271			Sequence 271, App
5	214	95.1	1025	61	US-60-118-318-159			Sequence 159, App
6	196	87.1	1097	25	US-60-172-360-2785			Sequence 2785, Ap
7	196	87.1	1097	25	US-09-649-163-7215			Sequence 7215, Ap
8	183	81.3	860	27	US-09-652-109-8606			Sequence 8606, Ap
9	172	76.4	1183	25	US-09-698-010-152			Sequence 152, App
10	172	76.4	1183	27	US-09-698-010-11426			Sequence 11426, A
11	45.2	20.1	1664976	13	PCT-US97-14900A-1			Sequence 1, Appli
12	45.2	20.1	1664976	13	US-08-916-421-1			Sequence 1, Appli
13	45.2	20.1	1664976	13	US-08-916-421B-1			Sequence 2534, Ap
14	42.2	18.8	538	81	US-60-377-240-2534			Sequence 3984, Ap
15	41.2	18.3	401	30	US-09-790-483-3984			Sequence 9112, Ap
16	40.8	18.1	397	22	US-09-553-094-9112			Sequence 173, App
17	40	17.8	6809	42	US-10-221-613-173			Sequence 7932, App
18	39.6	17.6	1443	22	US-09-577-410-7932			Sequence 279, App
19	39.4	17.5	88098	20	US-09-534-859-279			Sequence 279, App
20	39.4	17.5	88098	31	US-09-803-736-279			Sequence 3726, Ap
21	39.2	17.4	335	31	US-09-821-837-3726			


```
Db 61 ATATTGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120
Qy 121 TCCCCCAAAAGCAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCAT 180
Db 121 TCCCCCAAAAGCAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCAT 180
Qy 181 TTGATCTGCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225
Db 181 TTGATCTGCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225

RESULT 4
US-60-118-318-159
; Sequence 159, Application US/60118318
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 P
; CURRENT APPLICATION NUMBER: US/60/118,318
; CURRENT FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: PERL Program
; SEQ ID NO 159
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 4132547CT1
US-60-118-318-159

Query Match 100.0%; Score 225; DB 55; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGATGTTAGTTGTTTCAGGCATATAAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60
Db 1 TGATGTTAGTTGTTTCAGGCATATAAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60
Qy 61 ATATTGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120
Db 61 ATATTGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120
Qy 121 TCCCCCAAAAGCAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCAT 180
Db 121 TCCCCCAAAAGCAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCAT 180
Qy 181 TTGATCTGCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225
Db 181 TTGATCTGCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225

RESULT 5
US-60-172-360-2785
; Sequence 2785, Application US/60172360
; GENERAL INFORMATION:
; APPLICANT: Moris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; FILE REFERENCE: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
; FILE REFERENCE: GX-0007 P
; CURRENT APPLICATION NUMBER: US/60/172,360
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 29838
; SOFTWARE: PERL Program
; SEQ ID NO 2785
; LENGTH: 1025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 115406.1
US-60-172-360-2785

Query Match 95.1%; Score 214; DB 61; Length 1025;
Best Local Similarity 99.8%; Pred. No. 4.6e-36;
Matches 225; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TGATGTTAGTTGTTTCAGGCATATAAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60
Db 1 TGATGTTAGTTGTTTCAGGCATATAAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60
Qy 61 ATATTGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120
Db 61 ATATTGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120
Qy 121 TCCCCCAAAAGCAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCA 179
Db 121 TCCCCCAAAAGCAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCA 180
Qy 180 TTGATCTGCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225
Db 181 TTGATCTGCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225

RESULT 6
US-09-649-163-7215/c
; Sequence 7215, Application US/09649163
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Galvin, Katherine A.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kingsbury, Gillian A.
; APPLICANT: Welch, Nadine S.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Williamson, Mark
; APPLICANT: Richardson, Jennifer
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Villeval, Jean-Luc M.G.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; APPLICANT: Busfield, Samantha J.
; APPLICANT: Deeds, James
; APPLICANT: Lee, John
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1164-001
; CURRENT APPLICATION NUMBER: US/09/649,163
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/150,608
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 10535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7215
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1097)
; OTHER INFORMATION: n = A,T,C or G
US-09-649-163-7215

Query Match 87.1%; Score 196; DB 25; Length 1097;
Best Local Similarity 99.5%; Pred. No. 3.4e-32;
Matches 207; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 19 GCATATAAATTTGAAATAAATTTATGAGGCTCCATGATGCTATATTTGGTTTACCTTCA 78
```

Db 1092 GGCATATAATTTGAATAAATTAATGAGGCTCCATGATATGCTATATATGTTTACCTTCA 1033
Qy 79 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAA-GCTAGCTGTCCCAAAACGAAA 137
Db 1032 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAAAGGCTAGCTGTCCCAAAACGAAA 973
Qy 138 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 197
Db 972 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 913
Qy 198 CTGAATCAATTAGGAATTTTTTTTTTTT 225
Db 912 CTGAATCAATTAGGAATTTTTTTTTTTT 885

RESULT 7

US-09-652-109-8606/c
Query Match 87.1%; Score 196; DB 25; Length 1097;
Best Local Similarity 99.5%; Pred. No. 3.4e-32;
Matches 207; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
APPLICANT: McCarthy, Sean A.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.1180-001
CURRENT APPLICATION NUMBER: US/09/652,109
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/151,128
PRIOR FILING DATE: 1999-08-30
NUMBER OF SEQ ID NOS: 10105
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8606
LENGTH: 1097
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1097)
OTHER INFORMATION: n = A,T,C or G
US-09-652-109-8606

Qy 19 GGCATATAATTTGAATAAATTAATGAGGCTCCATGATATGCTATATGTTTACCTTCA 78
Db 1092 GGCATATAATTTGAATAAATTAATGAGGCTCCATGATATGCTATATGTTTACCTTCA 1033
Qy 79 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAA-GCTAGCTGTCCCAAAACGAAA 137
Db 1032 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAAAGGCTAGCTGTCCCAAAACGAAA 973
Qy 138 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 197
Db 972 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 913
Qy 198 CTGAATCAATTAGGAATTTTTTTTTTTT 225
Db 912 CTGAATCAATTAGGAATTTTTTTTTTTT 885

RESULT 8

US-09-698-010-152
Query Match 81.3%; Score 183; DB 27; Length 860;
Best Local Similarity 94.2%; Pred. No. 2e-29;
Matches 212; Conservative 0; Mismatches 10; Indels 3; Gaps 2;
APPLICANT: Williamson, Mark
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.2029-001
CURRENT APPLICATION NUMBER: US/09/698,010
CURRENT FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: 60/162,358
PRIOR FILING DATE: 1999-10-29

NUMBER OF SEQ ID NOS: 15684
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 152
LENGTH: 860
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(860)
OTHER INFORMATION: n = A,T,C or G
US-09-698-010-152

Qy 1 TGATGTGAAGTTGTTTTCAGGCATAAAATTTGAAATAAATATGAGCTCCATGATATGCT 60
Db 72 TGATGTGAAGTTGTTTTCAGGCATAAAATTTGAAATAAATATGAGCTCC-CATGATATG 130
Qy 61 ATATTGGTTTTCACCTTCAGAAGAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120
Db 131 CTTTGGTTTTCACCTTCAGAAGAATATTAGTTTCACTCAGGTTTTTCAAAG--TCGCTG 188
Qy 121 TCCCCCAAAAACGAAACAAAACAAAACAAACCTTTTAAAGAGTTGATGCTACTCAT 180
Db 189 TCCCCCAAAAACGAAACAAAACAAAACAAACCTTTTAAAGAGTTGATGCTACTCAT 248
Qy 181 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 225
Db 249 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 293

RESULT 9

US-09-649-163-9422
Query Match 76.4%; Score 172; DB 25; Length 1183;
Best Local Similarity 89.7%; Pred. No. 4.9e-27;
APPLICANT: Holtzman, Douglas A.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.1164-001
CURRENT APPLICATION NUMBER: US/09/649,163
CURRENT FILING DATE: 2000-08-25
PRIOR APPLICATION NUMBER: 60/150,608
PRIOR FILING DATE: 1999-08-25
NUMBER OF SEQ ID NOS: 10535
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9422
LENGTH: 1183
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1183)
OTHER INFORMATION: n = A,T,C or G
US-09-649-163-9422

LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
Query Match 20.1%; Score 45.2; DB 13; Length 1664976;
Best Local Similarity 51.0%; Pred. No. 24;
Matches 107; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 7 TAAGTTGTTTCAGGCATAAAATTTGAATAAATATGAGGCTCCATGATATGCTATATTG 66
DB 151403 TAAATCTTCGACCCCTTCTATTTCCATCCATTTATATAGCTACATGTTTGGTAAATTG 151462
QY 67 GTTTTACCTTCAGAGAATATTTAGTTTTCACCTCAGGTTTTTCAAAGCTAGCTGTCCCC 126
DB 151463 TCTTTATCAATAGCCTCTTCTAAATCTTCTCGTAAATTTTAAACTCCGGAGTTGCT 151522
QY 127 AAAAAAGGAAACAAACAAAAAACCTTTTAAAGTTGATGGCTACTCATTTTGATC 186
DB 151523 AAAAATCTATAAAATCAATTTACTCTTTTCCAAATCTTTTAAATTTGTTTTTATC 151582
QY 187 TGCCTCTCTGCTGAATCAATAGGAATTT 216
DB 151583 TTACCAAAATCCACCTATTAGGAATTT 151612
RESULT 14
US-60-377-240-2534/c
Sequence 2534, Application US/60377240
GENERAL INFORMATION:
APPLICANT: Diggans, James C.
APPLICANT: Porter, Mark
APPLICANT: Wei, Tao
TITLE OF INVENTION: Canine Gene Microarrays for Molecular Toxicology Modeling
FILE REFERENCE: 44921-5116-PR
CURRENT APPLICATION NUMBER: US/60/377,240
CURRENT FILING DATE: 2002-05-03
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 11109
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2534
LENGTH: 538
TYPE: DNA
ORGANISM: Canis familiaris
FEATURE:
OTHER INFORMATION:
US-60-377-240-2534
Query Match 18.8%; Score 42.2; DB 81; Length 538;
Best Local Similarity 55.0%; Pred. No. 33;
Matches 83; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 29 TTGAATAAATATGAGGCTCCATGATATGCTATATTTGGTTTTCCTTCAGAGAAATATT 88
DB 537 TTGACATACATTTTGATGTATGTTAGTTAGGCTATACAGAAATGTACAAACCAAAGCTGAT 478
QY 89 TAGTTTCACTCAGGTTTTTCAAAGCTACGCTGTCCCCCAAAACAAACAAACAAAA 148
DB 477 CAATAATACATAGGTTTTTCAAAAAATTAGGTTTCAATAAAGATACAGAGAAAGGAAAT 418

Qy 149 AACAACTTTTAAAGAGTGTGCTACTCA 179
Db 417 AACCAACATCCCAACACTCAGAACTTCTAA 387

RESULT 15

US-09-790-482-3984/c
; Sequence 3984, Application US/09790482
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2063-001
; CURRENT APPLICATION NUMBER: US/09/790,482
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/183,731
; NUMBER OF SEQ ID NOS: 4377
; SOFTWARE: Fast-SEQ for Windows Version 4.0
SEQ ID NO 3984
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(401)
; OTHER INFORMATION: n = A,T,C or G
US-09-790-482-3984

Query Match 18.3%; Score 41.2; DB 30; Length 401;
Best Local Similarity 55.5%; Pred. No. 51;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
Qy 23 TAAATTTGAAATTAATATGAGCTCCATGATATGCTATATGTTTACCTTCAAG 82
Db 154 TAAATGTCGAATCCCTACNTGACACATTATAACTGTATGGAGCTTAACCTTTATAAG 95
Qy 83 AATATTAGTTTCACTCAGCTTTTCAAGCTAGCTGTCCCAACAAACGAAACAAA 142
Db 94 GAAATGATTTTGACACTGGTATCTTTAAGATTTCTGATCCAAAAAAGCTTTGT 35
Qy 143 CAAAAAACAACCTTTT 159
Db 34 AAAAAAAGCTTTGT 18

Search completed: June 9, 2003, 09:40:39
Time: 1969 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:51 ; Search time 2538 seconds
(without alignments)
578.414 Million cell updates/sec

Title: US-09-765-231A-58
Perfect score: 225
Sequence: 1 ttagtgtagtggtttcagg.....attaggaattttttttttt 225

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7578120 seqs, 3262256734 residues

Total number of hits satisfying chosen parameters: 15156240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Pending Patents NA New.*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2.*
7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq3.*
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2.*
10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	100.0	320	9	US-10-313-542-271
2	225	100.0	320	9	US-10-313-542A-271
3	45.2	20.1	1664976	6	US-09-692-570-1
4	42.2	18.8	538	1	PCT-US03-13853-234
5	41.8	18.6	763446	10	US-60-470-166-8943
6	41	18.2	5518	9	US-10-240-453-180
7	39.6	17.6	31312	10	US-60-465-241-52707
8	39.6	17.6	31312	10	US-60-466-412-86559
9	39.2	17.4	60615	8	US-10-427-741-9
10	38.6	17.2	343317	10	US-60-465-241-51757
11	38.6	17.2	343317	10	US-60-466-412-84536
12	38.6	17.2	15105345	5	US-09-948-128-55
13	38	16.9	12393	9	PCT-US02-41414-1544
14	37.8	16.8	96588	1	US-09-837-604A-71071
15	37.6	16.7	109025	7	US-09-949-016-12609
16	37.6	16.7	109025	7	US-09-949-016-17567
17	37.6	16.7	2468502	7	US-09-947-911-346
18	37.6	16.7	2865598	7	US-09-947-911-26
19	37.6	16.6	1524	6	US-09-724-676-31394
20	37.4	16.6	1524	6	US-09-724-676A-31394
21	37.4	16.6	1524	6	US-09-724-676A-31394
22	37.4	16.6	2217	8	US-10-170-235-7647

ALIGNMENTS

RESULT 1
US-10-313-542-271
; Sequence 271, Application US/10313542
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/10/313,542
; PRIOR FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/495,050
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ IDS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 271
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4132547CT1
US-10-313-542-271
Query Match 100.0%; Score 225; DB 9; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.4e-42;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGAATGTAAGTGGTTTCAGGCATAAATTTGAATAAATATGAGGCTCCATGATGCT 60
DB 1 TGAATGTAAGTGGTTTCAGGCATAAATTTGAATAAATATGAGGCTCCATGATGCT 60
QY 61 ATATTGGTTTACCTTCAGAGAAATATTAGTTTCTACTCAGGTTTTTCAAGAGCTAGCTG 120
DB 61 ATATTGGTTTACCTTCAGAGAAATATTAGTTTCTACTCAGGTTTTTCAAGAGCTAGCTG 120
QY 121 TCCCCCAAAAAACGAAACAAAAACAAACAAACCTTTTAAAGAGTTGATGGTACTCAT 180
DB 121 TCCCCCAAAAAACGAAACAAAAACAAACAAACCTTTTAAAGAGTTGATGGTACTCAT 180
QY 181 TTGATCTGCTCTCTGCTGCTGAATCAATTAGGAATTTTTTTTTT 225
DB 181 TTGATCTGCTCTCTGCTGCTGAATCAATTAGGAATTTTTTTTTT 225

```
RESULT 2
US-10-313-542A-271
; Sequence 271, Application US/10313542A
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy M.
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice K.
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED
; PROTEINS
; FILE REFERENCE: PA-0013-1 CON
; CURRENT APPLICATION NUMBER: US/10/313,542A
; CURRENT FILING DATE: 2002-12-05
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
SEQ ID NO 271
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4132547CT1
US-10-313-542A-271

; Query Match 100.0%; Score 225; DB 9; Length 320;
; Best Local Similarity 100.0%; Pred. No. 2.4e-42;
; Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGATGGTAAGTTGTTTCAGGCATAAAATTTTGAATAAATTTATGAGGCTCCATGATGCT 60
Db 1 TGATGGTAAGTTGTTTCAGGCATAAAATTTTGAATAAATTTATGAGGCTCCATGATGCT 60

Qy 61 ATATTGGTTTTTACCTTCAGAGAAATTTTATGTTTCACTCAGGTTTTTCAAGCTACGCTG 120
Db 61 ATATTGGTTTTTACCTTCAGAGAAATTTTATGTTTCACTCAGGTTTTTCAAGCTACGCTG 120

Qy 121 TCCCCAAAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCAT 180
Db 121 TCCCCAAAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCAT 180

Qy 181 TTGATCTGCCTCTCTCTGCTGAATCAATAGGAATTTTTTTTTTTT 225
Db 181 TTGATCTGCCTCTCTCTGCTGAATCAATAGGAATTTTTTTTTTTT 225

RESULT 3
US-09-692-570-1
; Sequence 1, Application US/09692570
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
```

FEATURE:
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:

NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
Query Match 20.1%; Score 45.2; DB 6; Length 1664976;
Best Local Similarity 51.0%; Pred. No. 0.35; Mismatches 103; Indels 0; Gaps 0;
Matches 107; Conservative 0;
QY 7 TAAGTTGTTTCAGGCATAAAATTTGAAATAAATTTATGAGGCTCCATGATATGCTATATTG 66
Db 151403 TAAATCTCTGCAGCCCTTCTATTTCCATCCATTTATATAGTACATGTTTGGTAAATTG 151462
QY 67 GTTTTACCTTCAGAGAATATTAGTTTTCACTCAGSTTTTCAAAGCTAGCTGTCCCCC 126
Db 151463 TCTTTATCAATAGCCTCTCTTAAATCTTCTCGTAAATTTTAAAACTCCGGAGTTGTCT 151522
QY 127 AAAAAACGAAACAAACAAAAACAAACCTTTTAAAGAGTTGATGCTACTCATTTGATC 186
Db 151523 AAAATCTATAAAATCAATTTACTCTTTTCCAACTACTTTTAATTTGTTTTTATC 151582
QY 187 TGCCTCCTCTGCTGAATCAATTAGGAATTT 216
Db 151583 TTACCCAAATCCACCTATTAGGAATTT 151612

RESULT 4
PCT-US03-13853-2534/c
Sequence 2534, Application PC/TUS0313853
GENERAL INFORMATION:
APPLICANT: Diggins, James C.
APPLICANT: Porter, Mark
APPLICANT: Wei, Tao
TITLE OF INVENTION: Canine Gene Microarrays for Molecular Toxicology Modeling
FILE REFERENCE: 44921-5116-WO
CURRENT APPLICATION NUMBER: PCT/US03/13853
CURRENT FILING DATE: 2003-05-05
PRIOR APPLICATION NUMBER: US 60/377,240
PRIOR FILING DATE: 2002-05-03
NUMBER OF SEQ ID NOS: 11109
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2534
LENGTH: 538
TYPE: DNA

QY 7 TAAGTTGTTTTCAGGCATATAAATTGAAATAAATTATGAGCTCCATGATATGCTATATTG 66
Db 185157 TTATTTCTTTGAGGTAACAATTTCAAGAAATAATGTGATTTAATGTAATAATTTATAA 185216
QY 67 GTTTTACCTTCAGAGAAATATTAGT--TTCACTCAGGTTTTTCAAGAGCTACGCTGCTCC 124
Db 185217 ATCTGAAGACAAATGAATCTTTAGGAATCTACTCATTTATTTTAAAGATATGCTATCAG 185276
QY 125 CCAAAAAACGA 135
Db 185277 CCATATATCAA 185287
RESULT 12
US-09-948-128-55/c
; Sequence 55, Application US/09948128
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
; FILE REFERENCE: CL001294
; CURRENT APPLICATION NUMBER: US/09/948,128
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 465
; SEQ ID NO 55
; LENGTH: 15105345
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(15105345)
; OTHER INFORMATION: n = A,T,C or G
US-09-948-128-55
Query Match 17.2%; Score 38.6; DB 5; Length 15105345;
Best Local Similarity 65.9%; Pred. No. 11;
Matches 56; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 93 TTCACTCAGGTTTTTCAAGCTACGCTGCTCCCAAGAAACGAAACAAACAAACAAACAA 152
Db 6035217 TACAGCTGGGTAAACAGAGCGAACTGCTCTCAAAAAACAAACAAACAAACAA 6035158
QY 153 ACCTTTTAAAGATTGATGCTACT 177
Db 6035157 ATCTCTCAGAGTTCAGAGGAGCT 6035133
RESULT 13
US-10-311-455-1236/c
; Sequence 1236, Application US/10311455
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1236
; LENGTH: 12393
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:

; NAME/KEY: unsure
; LOCATION: 9300, 9362
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1236
Query Match 16.9%; Score 38; DB 9; Length 12393;
Best Local Similarity 49.5%; Pred. No. 17;
Matches 98; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 22 ATAAAAATTTGAAATAAATTTATGAGGCTCCATGATATGCTATATTGTTTACCTTTAGAA 81
Db 6387 AAAAAACCTTTAAAAAATAAATTTTAAAAATAATTTAAAAATAAATTTTCTTTTAAAA 6328
QY 82 GAATATTTAGTTTCTCAGGTTTTTCAAGCTACGCTGCTCCCAAAAAACGAAACAAA 141
Db 6327 AAAACATAATTTTAAACGATTTTAAAAAATAAATTTTAAAAATAAATTTTAAAAATAA 6268
QY 142 AAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATCTGCTCTCTGCTCA 201
Db 6267 TAAATAAAAAATCAAAATATACTCAATTTAACTAAATTTCTCCGCTCTCTCAATAACCTTA 6208
QY 202 ATCAATTAGGATTTTTT 219
Db 6207 AACAAATTTCTATCAATTT 6190
RESULT 14
PCT-US02-41414-1544/c
; Sequence 1544, Application PC/TUS0241414
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71249/RMS/DCF
; CURRENT APPLICATION NUMBER: PCT/US02/41414
; CURRENT FILING DATE: 2002-12-26
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1613
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1544
; LENGTH: 96588
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-41414-1544
Query Match 16.8%; Score 37.8; DB 1; Length 96588;
Best Local Similarity 51.5%; Pred. No. 18;
Matches 87; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 8 AAGTTGTTTCAGGCATATAAATTTGAAATAAATTTATGAGCTCCATGATATGCTATATTGG 67
Db 61750 ATGTTGAGCATTTGTATATTTTAAATTTTAAATTTTCCATGCTCAATTTATTAGTA 61691
QY 68 TTTTACCTTCAGAAGAATAATTTAGTTTCTCAGGTTTTTCAAGCTACGCTGCTCCCA 127
Db 61690 AATTTTAAATGAAGCATATATGATACATGATTTTGGAGAGCTGTTGTCAGCC 61631
QY 128 AAAAAACGAAACAAAAACAAACCTTTTAAAGAGTTGATGGCTAC 176
Db 61630 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 61582
RESULT 15
US-09-837-604A-71071/c
; Sequence 71071, Application US/09837604A
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: De La Pena, Robert C.

THIS PAGE BLANK (USPTO)